ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/280752164

Géométrie au XXe siècle

Article · January 2005

CITATION
1

READS
2,074

4 authors, including:

Flament Dominique
French National Centre for Scientific Research
32 PUBLICATIONS **77** CITATIONS
SEE PROFILE

Joseph Kouneiher
Université Côte d'Azur
193 PUBLICATIONS **1,003** CITATIONS
SEE PROFILE

All content following this page was uploaded by Joseph Kouneiher on 06 October 2015.

The user has requested enhancement of the downloaded file.

Géométrie au 20^{e} siècle

Histoire et Horizons

J. Kouneiher, D. Flament, P. Nabonnand, J.J. Szczeciniarz
éditeurs

Collection Actualités scientifiques

HERMANN EDITEURS DES SCIENCES ET DES ARTS

Préface

L'objectif de cet ouvrage est de livrer à l'ensemble de la communauté mathématique, scientifique ainsi qu'à un public plus large, un ensemble de travaux qui tentent de situer la place et d'expliquer le rôle que la géométrie comme discipline a pu jouer dans le développement du corpus des disciplines mathématiques. Nous avons aussi voulu en demandant aux auteurs de présenter leur vision historique et épistémologique de certains des sujets qui constituent la géométrie. Nous tenons à les remercier de la qualité des textes qu'ils ont produits. L'ensemble des textes ainsi livrés ne prétend, ce serait folie, nullement à l'exhaustivité mais il est néanmoins représentatif de l'activité mathématique. Il offre un matériau de travail pour les mathématiciens mais aussi pour les historiens et les philosophes des mathématiques désireux de s'aventurer sur le terrain contemporain.

L'histoire des mathématiques du vingtième siècle peut être considérée comme celle d'un moment d'apparition et d'autonomisation de nombreuses sous-disciplines au sein des mathématiques et en même temps d'interactions multiples et diverses entre les différents domaines des mathématiques.

Les derniers développements ont été de manière plus intenses que les précédents le lieu d'élaboration de théories qui représentent une puissante réflexion des mathématiques sur elles-mêmes, qui ont entraîné des niveaux complexes de stratification théorique et qui dans la plupart des cas affectent une grande partie du corpus mathématique. Depuis les *Principia Mathematica* et l'introduction de la théorie des ensembles au tournant du 20^{e} siècle, jusqu'au dernier théorème de Fermat, ce siècle a été le témoin de nombreuses tentatives d'unification partielles ou globales des disciplines mathématiques.

L'histoire des mathématiques est marquée dans ses moments essentiels par des tentatives de grande synthèse unifiant les directions parfois opposées suivant lesquelles elles se sont constituées. Ces grands pôles ont été et continuent d'être sous des formes plus complexes, à travers l'étroit entrelacement des disciplines, l'arithmétique et la géométrie. Il semble bien que la géométrie et devons-nous dire les géométries ont toujours exporté dans les autres disciplines plus centrées sur l'arithmétique ou l'algèbre leur mode de questionnement. L'exemple des mathématiques grecques est à cet égard symptomatique.

La grande tentative d'unification entre algèbre et géométrie qui a son point de départ essentiel avec l'œuvre de René Descartes s'est produite deux millénaires plus tard. Le programme cartésien a initié un type d'unification extrêmement fécond entre domaines des mathématiques. Ces formes d'interpénétration des domaines entre eux sont aujourd'hui courantes. Le programme cartésien est concrétisé par la naissance d'une nouvelle discipline, la géométrie analytique. Ce niveau et cette forme de liaison ont été atteints aujourd'hui quand

on considére l'analyse et la théorie des nombres.

Un autre élément caractéristique des développements ultérieurs des mathématiques est leur autonomisation à travers des formes spécifiques d'abstraction qui les voient accélérer les processus de réflexion sur elles-mêmes. Certaines branches des mathématiques s'engagent à des moments différents suivant des modalités diverses sur cette voie de "creusement de soi". La géométrie analytique via l'algèbre linéaire est devenue la théorie spectrale, la théorie des équations algébriques a donné naissance à la théorie des groupes et à l'algèbre générale moderne, la géométrie algébrique se trouve réinvestie par la réflexion du corpus algébrique qui en fait une véritable instrument d'exploration de la géométrie. Le processus d'abstraction dont nous n'avons encore que fort peu exploré les formes est un processus naturel et essentiel de la pensée mathématique. Car c'est aussi ce que nous voulons faire comprendre à savoir la théorie et la pratique mathématique sont des formes de pensée réelles et non pas seulement des constructions d'outils ou des jeux de langage ou de symboles. Yuri Manin écrit *"... les mathématiques associent à d'importantes abstractions physiques (modèles) ses propres constructions mentales, qui sont très éloignées des impressions directes des expériences et de l'expérimentation physique"*.

Le rôle de l'abstraction devient évident au cours de ce siècle. Selon Jean Dieudonné, les principaux intérêts des mathématiques depuis 1920 ont été d'étudier les structures plutôt que les objets. Cela apparaît dans l'importance accordée à des concepts tels que les invariants et les catégories. L'invention des structures va de pair avec une montée des constructions impliquant des objets de plus en plus abstraits et complexes, tandis que les structures deviennent à leur tour l' objet de la constitution de nouvelles structures. Pour autant, les géométries dans leur ensemble et la physique théorique moderne continuent d'entretenir des relations fécondes, illustrées à la fin du 20^{e} siècle, par le développement de théories dans lesquelles mathématiques et physiques sont indissociablement liées.[1]

Le but de cet ouvrage est donc de proposer une large revue synthétique des géométries du 20^{e} siècle ainsi qu'une présentation réflexive des fondements conceptuels de ces géométries. Les contributions qui suivent montrent que le questionnement géométrique traverse la plupart des domaines mathématiques (y compris le calcul numérique,[2] ce qui se traduit par l'émergence et le développement de nombreuses théories à caractère géométrique.

Géométrie différentielle

La géométrie différentielle joue un rôle central dans toutes les mathématiques. Au-delà de ses propres intérêts, elle fournit des idées et des techniques essentielles aux physiciens et ingénieurs et entretient des liens étroits avec des pans entiers de l'algèbre, de la topologie, de la géométrie non-euclidienne, de l'analyse en générale (avec, en particulier, la théorie des équations aux dérivées partielles), la théorie du contrôle et la mécanique.

Héritière, entre autres, des travaux de K. F. Gauss, B. Riemann, de F. Klein, de É. Cartan, la géométrie différentielle s'est développée au 20^{e} siècle en incluant des considérations locales telles que les invariants différentiels et les jets, aussi bien que des idées globales telles que la théorie de Morse et les classes caractéristiques. De plus elle se construit autour des

[1] Voir les contributions de M. Atiyah, J.-P. Bourguignon, A. Michel et E. Scholz.

[2] Voir la contribution de M. Zerner.

concepts de fonctions, de métriques, de fibrés, de dérivées, de lissage et de direction, et surtout de connexion et de courbure.[3]

Par exemple, le travail de Cartan[4] introduit la notion "d'espaces généralisés (géométries de Cartan)" généralisant les espaces homogènes (géométries de Klein) de la même façon que la géométrie de Riemann généralise la géométrie d'Euclide.

D'un autre point de vue, la théorie des jets apparaît comme une branche différente de la géométrie différentielle. Alors même que Ehresmann dans ses papiers originaux de 1951 insistait sur la signification conceptuelle de la notion de r-jet en géométrie différentielle, les jets étaient abondemment utilisés comme outil purement technique dans certains problèmes de la théorie des systèmes d'équations aux dérivées partielles, dans la théorie des singularités[5], dans le calcul variationnel et la mécanique d'ordre supérieur.

La théorie des feuilletages, née dans les années 1940 grâce aux travaux fondateurs d'Ehresmann et de Reeb connait à partir du résultat de Novikov dans les années 1970 un développement remarquable.[6] Au même moment, les conceptions de M. Gromov modifient radicalement la manière d'aborder les problèmes non seulement en géométrie riemannienne mais aussi en théorie des systèmes dynamiques et plus généralement dans presque tous les domaines de la géométrie.[7]

Le développement moderne de la géométrie différentielle globale met en avant la règle selon laquelle les objets de la géométrique différentielle forment des fibrés sur des variétés. Dans cette prespective des foncteurs peuvent être définis sur la catégorie des variétés fibrées et leurs isomorphismes locaux, le fibré des connexions générales en est un exemple simple.

D'autres exemples et applications peuvent être cités, tels que le théorème de Gilkey selon lequel toutes formes différentielles dépendants naturellement de métriques riemanniennes et satisfaisant certaines conditions d'homogénéité, sont en fait des formes de Pontryagin. C'est un résultat essentiel dans la démonstration du théorème de l'index d'Atiyah-Singer utilisant le "noyau de chaleur".

Enfin, depuis les travaux de Plateau sur les surfaces minimales[8], la géométrie différentielle a été la source de nombreux problèmes et d'inspiration pour la théorie des équations aux dérivées partielles. Ces développements récents mettent en jeu les applications harmoniques, des problèmes sur une courbure imposée, les équations de Monge-Ampère, les variétés de Kähler-Einstein, les invariants de Seiberg-Witten et leurs relations aux invariants de Gromov en géométrie symplectique.

Géométrie Symplectique

Les structures symplectiques apparaissaient principalement dans la formulation Hamiltonienne de la mécanique classique, et ce lien persiste depuis[9]. Même aujourd'hui des idées venant de la physique, de la théorie des cordes notamment, continuent à enrichir le développement de la géométrie symplectique. Les définitions de base de cette géométrie sont

[3] Voir les contributions de M. Berger et J.M. Souriau.
[4] Voir les contributions de P. Libermann et de A. Michel.
[5] Voir la contribution de M. Chaperon.
[6] Voir la contribution de A. Haefliger.
[7] Voir la contribution de A. Zeghib.
[8] Pour la version complexe, voir la contribution de P. Dolbeault.
[9] Voir les contributions de Ch.-M. Marle et A. Weinstein.

assez naturelles : on étudie la géométrie d'une 2-forme bilinéaire antisymétrique ω au lieu de celle d'une forme symétrique comme en géométrie riemannienne. Ce changement est radical. En effet, les mesures unidimensionnelles disparaissent du fait de l'antisymétrie de la forme. Ainsi la géométrie symplectique est essentiellement une géométrie bi-dimensionnelle qui mesure la surface d'une courbe complexe à la place de la longueur d'une courbe réelle. On constate dans cette géométrie le développement des thématiques suivantes :

1) L'étude des hypersurfaces et des sous-variétés lagrangiennes, qui apparaissent dans les questions de dynamique et l'étude des sous-variétés symplectiques qui sont étroitement liées à la géométrie riemanienne complexe, en particulier les courbes J-holomorphes que l'on peut considérer comme les analogues en géométrie symplectique des géodésiques de la géométrie riemannienne.

2) L'étude des questions liées à la nature topologique de cette théorie. En effet, bien que les concepts de cette géométrie s'expriment en termes de formes différentielles, ils ne dépendent pas intrinsèquement des dérivées.

3) Enfin, une des propriétés intéressantes des variétés symplectique est le lien *local-global* et l'existence d'un analogue global pour tout énoncé local. On peut remarquer à cet égard que l'analogue global d'un énoncé local linéaire peut être non linéaire, comme lors du passage de la géométrie de Kähler à celle symplectique. Remarquons ainsi que la distinction entre divers espaces symplectiques se fait au moyen des invariants globaux.

Géométrie Algébrique

La géométrie algébrique consiste en l'étude des structures et des formes géométriques provenant de l'analyse des structures algébriques des anneaux de polynômes. Au 20^{e} siècle on a découvert que les idées fondamentales de la géométrie algébrique classique pouvaient être appliquées à tout anneau commutatif contenant l'unité, tel que les entiers. La géométrie de ce type d'anneau est déterminée par sa structure algébrique, en particulier ses idéaux premiers. Alexandre Grothendieck définit en 1960 les *schémas* comme étant les objets géométriques de base, qui ont la même relation à la géométrie d'un anneau que celle d'une variété à une carte de coordonnées. Le langage de la théorie des catégories a évolué à peu près à la même époque, en réponse principalement au besoin croissant de l'abstraction en géométrie algébrique.[10]

En conséquence, l'utilité de la géométrie algébrique s'est accrue dans de nombreux autres domaines des mathématiques, notamment en théorie algébrique des nombres[11] ; ainsi, P. Deligne l'utilise pour prouver une variante de l'hypothèse de Riemann et la démonstration d'Andrew Wiles du dernier théorème de Fermat emploie des outils développés en géométrie algébrique. Dans la dernière partie du vingtième siècle, des mathématiciens ont essayé d'étendre la relation entre l'algèbre et la géométrie à des anneaux arbitraires non commutatifs. L'étude des géométries associées aux anneaux non-commutatifs est appelée géométrie non-commutative [12]. Au cours de cette période, le travail de Grothendieck a assuré l'unification de certains thèmes de la géométrie, la théorie des nombres, la topologie et l'analyse complexe. L'introduction de la théorie des schémas a permis la résolution de conjectures de

[10] Voir la contribution de F. Chargois.
[11] Voir l'article de C. Soulé.
[12] Voir la contribution de Y. Manin.

Weyl en théorie des nombres. Grothendieck donne également une définition algébrique du groupe fondamental d'une courbe.

D'un autre côté, le travail de Michael Atiyah a contribué à un grand nombre de sujets en mathématique centrés sur les interactions entre géométrie et analyse. Sa première grande contribution (en collaboration avec F. Hirzebruch) porta sur le développement d'une nouvelle technique puissante en topologie (K-théorie) afin de résoudre de nombreux problèmes difficiles restés en suspens. Par la suite, il établit (en collaboration de I. M. Singer) un important théorème concernant le nombre de solutions des équations différentielles elliptiques. Ce "théorème de l'indice" a des antécédents en géométrie algébrique et conduit à de nouveaux liens importants entre la géométrie différentielle, la topologie et l'analyse. Combiné à des considérations de symétrie, ce résultat induit (conjointement avec les travaux de R. Bott[13]) un nouveau "théorème du point fixe" plus élaboré.
De plus, le théorème de l'indice reçoit une interprétation en termes de théorie quantique et dans laquelle les théories de jauge mettent en jeu des équations différentielles non-linéaires aux propriétés profondes et intéressantes pour les mathématiciens. Un exemple est le résultat de Simon Donaldson en géométrie 4-dimensionnelle.

Comme dans d'autres domaines des mathématiques, la classification apparaît comme un souci majeur. La classification des variétés algébriques[14] a toujours constitué un problème fondamental en géométrie algébrique. Un progrès majeur dans la classification des surfaces algébriques a été obtenu dans la première partie du 20^{e} siècle par les grands géomètres algébristes italiens avec à leur tête Castelnuovo, Enriques et Severi. La contribution de Zariski dans les années 1950 et le travail de Kodaira au cours des 10 années suivantes s'inscrivent dans cette continuité. Le travail de Mori accomplit une remarquable continuation de ces efforts de classification en géométrie algébrique et en de nombreux points fournit un apport fort appréciable dans les progrès en géométrie algébrique au cours du vingtième siècle.

La géométrie algébrique réelle constitue un cas à part au sein de la géométrie algébrique qui, du fait de la spécificité de ses méthodes et de ses objets, tend à s'autonomiser. Ainsi, les nouvelles techniques de construction des variétés algébriques réelles de Viro ont permis des avancées significatives pour la classification des variétés algébriques réelles.[15] Un exemple surprenant d'interaction entre deux domaines qui pourraient *a priori* paraître assez éloignés l'un de l'autre est celui de la géométrie algébrique réelle et la théorie des modèles qui ont un résultat fondateur commun avec le travail de Tarski consacré aux ensembles semi-algébriques.[16]

Topologie algébrique

La topologie algébrique se constitue en théorie autonome à la fin du 19^{e} siècle avec en particulier le travail de Henri Poincaré. Si, actuellement, les théories d'homologie, de cohomologie ou d'homotopie sont autant d'outils fondamentaux pour l'étude des géométries, une des questions essentielles de la topologie algébrique, avant la mise au point de ces

[13] Voir la contribution de R. Bott.
[14] Voir la contribution de F. Campana.
[15] Voir la contribution de I. Itenberg.
[16] Voir la contribution de M.-F. Roy.

techniques à caractère algébrique, était de classer les variétés aux homéomorphies près.[17] La conjecture de Poincaré, un des plus célèbres problèmes mathématiques du 20^{e} siècle, vient, semble-t-il, d'être résolue positivement.[18] Celle de Thurston n'est pas encore résolue.[19]
Ainsi, au début, l'objectif de la topologie algébrique était d'étudier des objets géométriques tels que les variétés lisses, qui appparaîssent en rapport avec les équations différentielles en physique, en transformant des géométries de dimensions supérieures en des algèbres plus accessibles à l'aide de diverses théories cohomologiques. La théorie d'homotopie stable est le cadre ultime qui nous permet de réaliser ce type de conversion des données géométriques à celles algébriques. Les théories de cohomologies sont elles-mêmes représentées par des objets géométriques et c'est seulement très récemment que des modèles satisfaisant ont été obtenus, ouvrant un champ vaste de nouvelles possibilités pour exploiter les idées algébriques en topologie, géométrie, théorie des nombres et physique. La topologie algébrique a été utilisée avec succès par Lefschetz, Hodge et d'autres pour aborder de nombreux problèmes en géométrie algébrique complexe. Le développement des théories cohomologiques qui, dans le cas de la caractéristique positive, se comportent comme la cohomologie classique utilisée par Lefschetz a été un des résultats majeur du 20$^{\text{ième}}$ siècle.

Géométrie complexe

L'analyse complexe et la géométrie analytique vont de pair et ont de nombreuses applications à la fois dans d'autres domaines des mathématiques pures (géométrie algébrique, géométrie différentielle, systèmes dynamiques, E.D.P., topologie, théorie de nombres, etc.) et en mathématiques appliquées (physiques théoriques, géophysiques, mathématiques économiques, tomographie, ...).

L'imbrication entre géométrie et analyse complexe est bien illustrée par l'histoire du problème de Plateau complexe qui trouve son origine dans des considérations spécifiquement analytiques comme la question de l'extension d'une fonction en une fonction holomorphe, sa formulation la plus claire en termes de variétés minimales et les outils de sa résolution grâce à la théorie géométrique de la mesure et la théorie des équations aux dérivées partielles.[20]

Une accélération du développement de ce domaine est sensible dans les années 1965, lorsque plusieurs théories quantitatives puissantes des équations de Cauchy-Riemann ont été développés (théories-L^2 de L. Hörmander et J. J. Kohn/ L. Nirenberg, travaux H. Grauert, G. Henkin et d'autres sur les noyaux de Cauchy-Fantappié) et que le concept de courants plurisousharmoniques a été correctement établi[21]. Le développement et l'élaboration de ces théories ont dominé la période entre 1965 et 1989, ainsi que la théorie des résidus et le fondement des CR–variétés. Ces nouvelles méthodes ont rendu possible l'étude des liens entre la géométrie et l'analyse.
D'un autre côté, les fibrés vectoriels et les méthodes transcendantes au coeur de l'analyse complexe et de la géométrie ont été utilisés en théorie de twisteurs et en analyse non linéaire. La théorie de Cauchy-Riemann quand elle est utilisée dans la résolution des problèmes de

[17] Voir la contribution de K. Volkert.
[18] Voir la contribution de V. Poenaru.
[19] Voir la contribution de M. Boileau.
[20] Voir la contribution de P. Dolbeault.
[21] Voir la contribution de P. Lelong.

E.D.P. Les résidus interviennent dans la théorie des intégrales de Feynman, en géophysique et dans la résolution des équations de physique mathématique.

Géométrie Non-commutative

La géométrie non-commutative avait à l'origine pour but de géométriser des anneaux non-commutatifs arbitraires en associant à un anneau non commutatif un "spectre non-commutatif" : on étend la construction de spectres premiers d'un anneau commutatif et ensuite on "recolle" ces spectres en un "schèmas non-commutatif". Cependant le problème de recollement s'est avéré être très difficile et sa résolution ne se conforme pas aux attentes initiales. La raison en est que les spectres non-commutatifs généraux ne sont pas suffisamment fonctoriels pour permettre leurs localisations (ou leurs recollements) de manière usuelle.
Néanmoins, la géométrie non-commutative a fait des progrès considérables, et depuis les années 1970, l'idée même de la géométrie non-commutative est devenue plus élaborée et plus complexe, ce qui explique le fait que cette géométrie est devenue un des sujets le plus étudié en mathématiques et en physique mathématiques durant la dernière décennie.
Elle est aujourd'hui un des sujets les plus fascinants, ne connaissant aucune limite interdisciplinaire et où cohabitent géométrie algébrique de caractéristique p et intégrales de Feynman.

Plus récemment, partant de l'analyse fonctionnelle Alain Connes a induit un développement majeur qui a conduit à redéfinir globalement le sujet. Il s'agit de "géométriser" des C^*–algèbres. Ce qui est crucial est le fait qu'en passant des modèles locaux, Alain Connes a réussi à construire directement, pour ces espace non-commutatifs, des invariants type K–théorie analogues à ceux de la géométrie différentielle. L'outil essentiel pour cette construction est l'analogue non-commutatif du complexe de de Rham.

La construction de groupes quantiques en 1982 par Vladimir Drinfeld a représenté un autre évènement marquant en géométrie non-commutative dans les années 1980. Drinfeld a decouvert que les travaux de Ludwig Faddeev et son école qui traitent des solutions des équations de Yang-Baxter, impliquent des algèbres de Hopf non-commutatives (et non-cocommutatives) qui sont des déformations des algèbres de fonctions sur des groupes de Lie classiques. Ces nouvelles algèbres de Hopf sont " les anneaux de fonctions de groupes quantiques". Ces groupes (quantiques) et leurs représentations sont devenus en un temps très court une nouvelle branche des mathématiques, qui s'est développée et porte aujourd'hui l'appellation géométrie quantique.
Mais, comme en géométrie non-commutative de Connes, le point central de cette géométrie quantique s'est écarté de l'idée de la géométrisation d'origine, et l'accent est mis sur la quantification d'une géométrie commutative, i.e., on déforme d'une manière convenable les anneaux de fonctions usuelles ainsi que les faisceaux qui sont au dessus de ces anneaux, puis on traîte ces objets déformés comme s'ils étaient des fonctions et des faisceaux sur des "espaces quantiques". Un exemple intéressant qui marie la géométrie quantique et le style de la géométrie non-commutative de Connes est le tore quantique et les fonctions theta quantifiées de Manin[22].

[22]Voir les contributions de Y. Manin.

Mathématisation, Physique Mathématique et géométrisation

La mathématisation de la science n'a longtemps porté que sur les domaines de l'astronomie (Ptolémée) et de la statique (Archimède). C'est seulement au 17^{e} siècle que se sont constitués les concepts de base qui président à l'élaboration d'une physique mathématique. Dans les autres domaines – notamment l'optique, la chimie, la physiologie – on rencontre certes, même avant le 17^{e} siècle, certaines formes de mathématisation, limitées le plus souvent à celles d'une quantification qui ne se structure pas comme telle (même si elle peut prendre parfois des formes subtiles) ; la géométrisation en optique n'est envisagée que du seul point de vue de la propagation des rayons lumineux.

Il faut attendre la fin de 18^{e} siècle pour voir des domaines comme l'électricité et le magnétisme se mathématiser en partie. Le 19^{e} siècle verra non seulement l'extension de la mathématisation à de nombreux autres domaines, surtout la chaleur, l'optique, l'électromagnétisme, mais aussi la mise en œuvre de cette mathématisation sous un mode beaucoup plus systématique, avec l'élaboration d'une physique mathématique.

Les mathématiques et la physique ont toujours eu des rapports très étroits. Mais, le développement de la connaissance et la spécialisation croissante des scientifiques à la fin du dix-neuvième siècle entraînent une séparation plus nette des deux disciplines. Bien que continuant à s'influencer mutuellement, les deux disciplines semblent diverger effectivement au milieu du 20^{e} siècle. Dans la dernière partie de ce siècle, une réunification remarquable a lieu, centrée principalement autour du contenu mathématique de la théorie quantique des champ et de la théorie des cordes[23]. L'exemple de la théorie des spineurs, introduite en 1913 par É. Cartan, utilisée et développée par des physiciens théoriciens comme Dirac et redécouverte dans la seconde moitié du 20^{e} siècle par les mathématiciens[24] illustre bien à la fois la permanence des influences mutuelles entre les deux domaines mais aussi l'évolution de leurs rapports.

La géométrie, qui a toujours constitué une partie essentielle des mathématiques, a pourtant longtemps été considérée comme la science d'une entité "naturelle" sensible, l'espace. Elle est même conçue par Pascal comme une véritable science expérimentale[25]. La question de la naturalité de l'espace qui ferait de la géométrie une science davantage tournée vers l'expérience est une question complexe. Il faut distinguer, au delà de leurs chevauchements et de leurs intrications, les idéalités empiriques morphologiques et les objets de la géométrie. Si la géométrie a été immergée depuis ses débuts dans l'univers des formes perçues et si cette immersion a renforcé une certaine conception de sa naturalité largement étayée par la conception de la figuration dans la géométrie euclidienne, il est vrai que l'axiomatisation proposée par les *Grundlagen der Geometrie* de Hilbert a permis de poser le problème d'une tout autre façon. Elle a pour son temps libéré les idéalités géométriques des liens qui les bridaient en les enchaînant aux idéalités morphologiques. Il est vrai qu'elle permet de reposer à d'autres frais la question de l'origine de nos concepts géométriques en montrant que situer la source de ceux-ci dans les propriétés des objets du monde, dont il faut les abstraire,

[23] Voir les contributions de M. Atiyah et de Y. Manin.
[24] Voir les contributions de J.-P. Bourguignon et de E. Scholz.
[25] Voir la contribution de G. Heinzmann.

reste largement prisonnier d'une conception naïve du travail mathématique. Ce qu'a montré l'histoire de la géométrie c'est qu'à travers la thématisation de l'objet surface comme objet géométrique doué de propriétés intrinsèques, puis du concept de variété comme concept susceptible de réalisations diverses et doué de propriétés géométriques, qu'il est alors possible de faire apparaître et d'étudier comme les propriétés métriques, topologiques, de courbure etc. c'est d'abord que l'espace lui-même est à la fois un objet formel relevant de la variété et un objet qui possède une certaine espèce de réalité susceptible d'interprétation physique.

La naturalité de la géométrie et donc l'originalité du travail de géométrisation doit être plutôt vue comme une élaboration incessante à partir des formes d'appréhension de l'abstrait qui passe par des modalités de visualisation dont nous sommes encore loin d'avoir percé les secrets et qui de fait procède à des déréalisations de certaines formes perceptives, mais par ailleurs parvient à en hisser d'autres à des niveaux de stratification complexe. Certains diront que la géométrie a disparu en passant des courbes aux variétés puis aux schémas, d'autres qu'elle s'est profondément transformée et qu'elle continue d'agir au sein même des procédés de faisceautisation. Il semble admis et reçu par l'ana- lyse que si les mathématiques manifestent cette omniprésence de l'intuition dont parle Hermann Weyl (*Philosophie der Mathematik und Naturwissenschaft* (1928)), l'intuition géométrique doit être conçue à travers les diverses formes de son déplacement le long des structures qui la transportent.

Nous voudrions ajouter deux remarques : d'abord le fait de hisser l'intui- tion le long du formalisme, qui s'en trouve souvent réadapté ou retravaillé en retour, nous indique que la géométrisation est également un processus hau tement opératoire, que l'intuition possède elle aussi des propriétés qui la font opérer, bref qu'elle est active ; ensuite le type de pensée qui caractérise la pensée mathématicienne, comme pensée opératoire provient pour une grande part de ce qu'elle est à même de prendre appui à son tour sur les actes les plus purs de la géométrie.[26]

Remerciements

Les contributions à ce livre sont issues du colloque *Géométrie au 20^e siècle : 1930-2000* qui s'est déroulé du 24 au 29 septembre 2001 à Paris. Nous tenons à remercier la Maison des Sciences de l'Homme de Paris, l'IHÉS, l'Institut Henri Poincaré, l'Académie des Sciences, la Société Mathématique de France, les universités Paris 6, Paris 7, Nancy 2, l'École normale supérieure de la rue d'Ulm, le CNRS et la Fondation Scott l'Erigène qui ont permis par leur soutien institutionnel, financier ou logistique le bon déroulement de ce colloque. Nos remerciements vont aussi au directeur de l'IHÉS, Jean-Pierre Bourguignon, sans le soutien duquel ce colloque n'aurait pu avoir lieu ainsi qu'aux membres du conseil scientifique (Alain Chenciner, Christian Houzel, Bernard Teissier, Jean Petitot et Marc Chaperon) qui n'ont pas ménagé leur peine pour assurer le succès du colloque et la parution de ce livre.

[26] Voir la contribution de G. Longo.

Table des matières

Préface ··· **iii**

Geometry and Physics of the 20th Century ··· *M. Atiyah* **3**
1 Historical and Philosophical Remarks . . . 3
2 Space and Time . . . 3
3 Probing the Unknown . . . 4
4 Scales . . . 5
5 Mathematics ahead of Physics . . . 5
6 Simultaneous Discovery . . . 6
7 Physics ahead of Mathematics . . . 7
8 M-theory . . . 9
9 Relation of Mathematics to Physics ? . . . 9

On The Early History of the Periodicity Theorem ··· *R. Bott* **11**

Von Zahlen und Figuren ··· *Yu. I. Manin* **31**
1 Introduction . . . 31
2 Real multiplication : an introduction . . . 33
2.1 Case (KW). . . . 34
2.2 Case (CM). . . . 34
2.3 Case (D). . . . 35
2.4 Case (LT). . . . 36
2.5 Real multiplication (RM) and noncommutative geometry. . . . 36
3 Noncommutative geometry and arithmetic . . . 37
3.1 Jones indices. . . . 37
3.2 Bost–Connes spontaneous symmetry breaking. . . . 38
3.3 $Gal\,(\overline{\mathbb{Q}}/\mathbb{Q})$ and the Grothendieck–Teichmüller group. . . . 39
3.4 Summary. . . . 39
4 Real multiplication of quantum tori : geometry . . . 40
4.1 Lattices and pseudolattices. . . . 40
4.2 Morita category and projective modules. . . . 41
4.3 Morita category of two–dimensional quantum tori. . . . 43
4.4 Complex Multiplication and Real Multiplication. . . . 43

4.5 Quantum tori as "limits" of elliptic curves. 44
5 Stark's numbers for real quadratic fields 46
5.1 Stark's numbers at $s = 0$. 46
5.2 Stark's conjecture for real quadratic fields. 47
5.3 Hecke's formulas. 47
5.4 Theta functions of pseudolattices. 47
5.5 Proposition 49
5.6 Poisson formula 50
5.7 Functional equations for θ and Θ 50
5.8 Proposition 52

La géométrie des nombres ··· *C. Soulé* **57**

Géométrie complexe et positivité ··· *P. Lelong* **67**
1 Introduction 67
2 Débuts et difficultés d'une géométrie complexe 68
3 La positivité dans les espaces complexes 69
3.1 Origine et bases du concept 69
3.2 Formes décomposables 70
3.3 Courants positifs 71
4 Fonctions plurisousharmoniques (Psh) et positivité 71
4.1 Propriétés générales et singularités impropres 71
4.2 Pseudo-convexité et domaines d'holomorphie 73
5 Courant positif fermé d'intégration 74
6 Géométrie des courants positifs fermés 75
6.1 Mesures positives et nombres (L) 75
6.2 Géométrie des nombres L et théorème de Siu 77
6.3 Propriétés de prolongement et d'approximation 78
7 Les courants positifs et les géométries complexes 79
7.1 Produits de courants positifs fermés 79
7.2 Opérateur de Monge-Ampère et géométrie pluricomplexe 80
7.3 Une notion de capacité pour la géométrie complexe 81
7.4 Une généralisation arithmétique 81

Sur le problème de Plateau complexe ··· *P. Dolbeault* **85**
1 Introduction 85
2 Origine possible : extension holomorphe ; enveloppe polynômiale d'une courbe réelle 85
3 Minimalité des ensembles analytiques complexes dans une variété kählérienne 86
3.1 Inégalité de Wirtinger (1936) 86
3.2 Variété kählérienne 87
4 Problème du bord comme problème d'extension géométrique : problème de Plateau dans le cas kählérien 88
5 Solutions du problème du bord dans différents espaces 89
6 Problème du bord et géométrie symplectique 94

L'influence de la théorie de Galois sur l'œuvre de Grothendieck ··· *F. Chargois* **97**
1 Introduction . . . 97
2 La théorie de Galois classique et la reformulation de Grothendieck . . . 99
2.1 Rappels sur la théorie de Galois . . . 99
2.2 Analogies avec la théorie topologique des revêtements . . . 99
2.3 La théorie de Galois revue par Grothendieck . . . 100
2.4 Bref retour sur la théorie topologique . . . 101
2.5 Quelques conséquences des idées de Grothendieck . . . 101
3 Catégories galoisiennes et groupe fondamental étale des schémas . . . 103
3.1 Catégories et dualité galoisiennes . . . 103
3.2 Groupe fondamental étale des schémas . . . 104
3.3 Le cas d'un schéma normal . . . 105
3.4 La suite exacte fondamentale d'homotopie . . . 105
3.5 Le théorème d'existence de Riemann . . . 106
4 Catégories tannakiennes et motifs . . . 106
4.1 Introduction . . . 106
4.2 Les origines de la théorie . . . 107
4.3 L'explication motivique et le lien avec Galois . . . 108
4.4 Catégories et dualité tannakiennes . . . 109
4.5 Retour aux motifs . . . 111
4.6 Cycles algébriques et retour à Galois . . . 112
4.7 Pour conclure . . . 112
5 Géométrie anabélienne, théorie de Galois-Teichmüller et dessins d'enfant . . . 113
5.1 Introduction . . . 113
5.2 Géométrie algébrique anabélienne . . . 113
5.3 Théorie de Galois-Teichmüller et dessins d'enfant . . . 114
5.4 L'apothéose de la théorie de Galois . . . 115

Les programmes de classification des variétés algébriques complexes : 1970-2000 ··· *F. Campana* **117**
6 Introduction . . . 117
7 Les variétés projectives complexes . . . 118
8 Les Courbes . . . 119
9 Le signe du fibré canonique . . . 120
10 Les Surfaces . . . 122
11 Le programme d'Iitaka-Moishezon . . . 124
12 1980- 2000 : La dimension trois, les courbes rationnelles et le programme des modèles minimaux . . . 125
13 Remerciements . . . 133

Géométries en dimension trois : de H. Seifert à W. Thurston ··· *M. Boileau* **139**
1 Introduction . . . 139
1.1 Les huit géométries homogènes en dimension trois . . . 141
1.2 Des variétés fibrées de Seifert aux variétés graphées de Waldhausen . . . 144

1.3 Les variétés hakeniennes . . . 146
1.4 De Kneser à Jaco, Shalen et Johannson . . . 147
1.5 Le Théorème d'hyperbolisation de Thurston . . . 148
1.6 Le cas fibré sur le cercle . . . 151
1.7 Le cas non fibré . . . 151
1.8 Quelques résultats dans le cas non hakenien . . . 152

Autour de l'hypothèse de Poincaré ··· *V. Poénaru* **161**
1 Introduction . . . 161
2 Brève histoire de l'hypothèse de Poincaré . . . 162
3 Un très bref aperçu sur le programme de Thurston . . . 167
4 Le programme de Po pour l'hypothèse de Poincaré . . . 167
5 La conjecture d'uniformisation topologique et la simple connexité à l'infini des revêtements universels (en dimension $n = 3$) . . . 179
6 Mathématiques et ontologie (Un essai provisoire) . . . 184

Le retour de la géométrie ··· *K. Volkert* **191**
1 La situation de la topologie vers la fin des années 1920 . . . 191
2 Les premiers travaux de Seifert et Threlfall . . . 193
3 Le problème algébrique . . . 195
4 Le problème géométrique : le domaine fondamental . . . 196
5 La théorie de Seifert . . . 200
6 Quelques remarques comme conclusion . . . 202

Histoire des immersions isométriques ··· *A. Zeghib* **209**
1 Introduction . . . 209
2 Début de la géométrie extrinsèque . . . 210
3 Naissance de la géométrie non-euclidienne . . . 210
4 Géométrie riemannienne . . . 211
5 Modèles de géométrie hyperbolique . . . 211
6 Problème d'immersions isométriques . . . 211
7 Obstructions : Théorème de Hilbert . . . 212
8 Courbure positive, cas des surfaces : rigidité et existence . . . 212
9 Obstructions. cas générique . . . 213
10 Cas analytique : Théorème de Cartan-Janet, systèmes différentiels, et autres approches . . . 213
11 Cas C^1 : Travaux de Nash (1954) et Kuiper (1955) . . . 214
12 h-principe de Gromov . . . 214
13 Cas C^k : Grand Théorème de Nash . . . 215
13.1 Amélioration de $q(n)$ . . . 215
13.2 Cas C^2 . . . 215
13.3 Le h-principe de Gromov . . . 215
14 Théorème de la fonction implicite dans les espaces de Fréchet . . . 216
15 Cas pseudo-riemannien . . . 216
16 Obstructions en codimension relativement petite . . . 217

16.1 Géodésibilité d'une sous-variété . . . 217
16.2 Lemme algébrique . . . 218
16.3 Le feuilletage local $\mathcal{F}l$ . . . 218
16.4 Intervalles . . . 218
17 Coup de théâtre : Preuve de Günther . . . 219

150 ans de Géométrie Riemannienne ··· *M. Berger* **223**
1 Introduction . . . 223
2 L'héritage de Riemann en langage moderne . . . 224
3 Formes d'espace, espaces localement symétriques (Riemann plus Klein) . . . 226
3.1 Espaces 3-transitifs . . . 226
3.2 Espaces 2-transitifs . . . 226
3.3 Espaces symétriques (localement) . . . 227
3.4 Espaces homogènes . . . 227
4 Programme de H. Hopf I : courbure et topologie . . . 228
4.1 Introduction . . . 228
4.2 Courbure presque constante . . . 228
4.3 Finitude et compacité . . . 229
4.4 Courbure de signe constant . . . 230
4.5 Courbure de Ricci. . . . 231
4.6 Courbure scalaire. . . . 232
5 Programme de Hopf II : métrisation, meilleure métrique : . . . 232
5.1 Métriques critiques pour une fonctionnelle . . . 233
5.2 Le miracle kählérien . . . 234
5.3 Pas d'espoir d'autre miracle . . . 234
6 Variétés riemannienne comme systèmes dynamiques . . . 235
6.1 Géodésiques périodiques . . . 235
6.2 Le flot géodésique. . . . 236
7 Le laplacien sur les fonctions : spectre, fonctions propres . . . 237
8 Quelles généralisations ? . . . 239

La géométrie différentielle d'Élie Cartan à Charles Ehresmann et André Lichnerowicz ··· *P. Libermann* **247**
1 Introduction . . . 247
2 Première partie : sur les travaux d'Élie Cartan (1869-1951) . . . 248
2.1 Systèmes différentiels et calcul différentiel extérieur . . . 248
2.2 Groupes continus "infinis" . . . 249
2.3 Géométrie Différentielle et ses applications à la Mécanique Analytique et à la Relativité . . . 250
3 Deuxième partie : Sur les travaux de C. Ehresmann (1905-1979) . . . 253
3.1 Espaces fibrés . . . 253
3.2 Connexions infinitésimales et connexions "de Cartan" . . . 254
3.3 Pseudogoupes, jets, groupoïdes de Lie . . . 255
3.4 Remarques . . . 259
4 Troisième partie : Sur certains travaux de A. Lichnerowicz (1915-1998) . . . 260

4.1 Sur les structures de Poisson 261
4.2 Sur les stuctures de Jacobi 263
4.3 Sur l'enseignement d'A. Lichnerowicz 263
5 Conclusion 264

La géométrie et le principe d'idonéité : une relecture de Ferdinand Gonseth ··· *G. Heinzmann* **269**
1 Qu'est-ce que la philosophie de la géométrie/des mathématiques ? 269
2 La solution de Gonseth 272

Espace et temps physiques et description des systèmes mécaniques ··· *C. M. Marle* **279**
1 Introduction 279
2 Espace et temps absolus de la mécanique newtonienne 280
2.1 Temps absolu 280
2.2 Espace absolu 282
3 Temps absolu, caractère relatif de tout mouvement 283
3.1 L'idée du caractère relatif de tout mouvement 283
3.2 L'espace-temps fibré sur le temps et sa structure affine 286
3.3 Le groupe de Galilée 288
3.4 Descriptions intrinsèques de l'évolution d'un système mécanique . . . 288
4 Abandon du temps absolu : la Relativité restreinte 291
4.1 Vitesse de la lumière, expériences de Michelson et Morley 291
4.2 L'espace-temps de Minkowski 293
4.3 Le groupe de Poincaré 294
5 La Relativité générale 294
5.1 Insuffisances de la Relativité restreinte 294
5.2 Espace-temps et équations d'Einstein 295
6 Remerciements et conclusion 296
7 Appendice 297

The Geometry of Momentum ··· *A. Weinstein* **307**
1 Introduction 307
2 Momentum in the calculus of variations 308
3 Hamiltonian systems with symmetry 309
4 Symplectic groupoid actions 311
5 Poisson Lie groups and beyond 312
6 Universal momentum maps 313
7 The Baer groupoid 315

Jets, transversalité, singularités : petite introduction aux grandes idées de René Thom ··· *M. Chaperon* **319**
1 Position du problème 319
2 Étude géométrique de $S = f^{-1}(0)$ 320
2.1 Génériquement, S est une surface C^∞ 320

2.2 Génériquement, Σ est une courbe C^∞ . . . 322
2.3 Formes normales . . . 323
3 "Digression" au cœur du sujet . . . 325
3.1 Conflits et bifurcations . . . 325
3.2 Multijets . . . 327
3.3 La stratification naturelle des espaces de fonctions . . . 327
4 Retour à l'équation (1.9) . . . 328
4.1 La stratégie de Lie . . . 329
4.2 La stratégie de Gromov . . . 330
5 Appendice : preuve de la proposition 3.1 . . . 330

Naissance des feuilletages, D'Ehresmann-Reeb à Novikov ··· *A. Haefliger* **335**
1 Préliminaires . . . 335
2 La thèse de Reeb . . . 337
3 Compléments dûs à Ehresmann . . . 341
4 De Reeb à Novikov . . . 343
5 Quelques problèmes issus de la thèse de Reeb . . . 346

La redécouverte des spineurs par les mathématiciens dans la seconde moitié du XXème siècle ··· *J. P. Bourguignon* **355**
1 Les spineurs avant 1930 . . . 355
2 Des années 30 aux années 50 chez les mathématiciens . . . 357
3 Le tournant du théorème de l'indice . . . 358
3.1 L'énoncé . . . 358
3.2 L'apparition de l'opérateur de Dirac et des spineurs . . . 358
3.3 Premières applications . . . 359
4 Les à-côtés topologiques . . . 360
5 Les à-côtés géométriques . . . 361
5.1 Spineurs harmoniques et existence de métriques à courbure scalaire positive . . . 361
5.2 Une autre approche du théorème de la masse positive en Relativité Générale . . . 363
5.3 L'approche spinorielle de l'espace-temps de Roger Penrose et les twisteurs . . . 363
5.4 L'approche par l'équation de la chaleur et le théorème de l'indice des familles . . . 364
6 La supersymétrie et la supergéométrie . . . 364

Local spinor structures in V. Fock's and H. Weyl's work on the Dirac equation (1929) ··· *E. Scholz* **371**
1 Introduction . . . 371
2 Setting the stage in the later 1920's for Weyl and Fock . . . 373
3 Weyl's and Fock's local spinor structure . . . 376
4 An additional $U(1)$-gauge . . . 381
5 Geometry and physics : interpretations and perspectives . . . 385

Les groupes comme Universaux ··· *J. M. Souriau* **395**

Géométrisation de la théorie physique : sur la genèse d'un problème ··· *A. Michel* **417**

1 De Kant à Riemann : l'horizon philosophique du problème. 418
1.1 Le problème de la géométrisation comme problème transcendantal. . . 418
1.2 La position riemannienne du " problème de la géométrisation " 422
1.3 Nature philosophique des prises de position quant au " problème de l'espace " 423
2 L'algébrisation du problème : l'influence de F. Klein et du " Programme d'Erlangen " 424
3 Le rôle de la relativité : la synthèse géométrique d'E. Cartan 428
3.1 La leçon de la relativité 428
3.2 Structure de la synthèse d'E. Cartan 430
3.3 Signification historique de la généralisation d'E. Cartan 433
4 Géométrisation de la physique et réflexivité 434

Topologie et cosmologie ··· *J. M. Luminet* **439**

Construction of real algebraic varieties ··· *I. Itenberg* **453**

1 Hilbert's 16-th problem 453
2 Combinatorial patchworking 458
3 Applications of combinatorial patchworking 460
3.1 Counterexamples to Ragsdale conjecture 460
3.2 Betti numbers of real algebraic surfaces 463

Some Topologies for Computations ··· *G. Longo* **473**

1 Introduction 473
1.1 Computations 473
1.2 Dimensions 475
2 Scott Topologies and Computable Functionals 476
2.1 Continuous Domains 477
2.2 Myhill-Shepherdson Theorem 478
2.3 Total Functionals, Filter Spaces and Real Numbers 481
3 Types 482
4 The Hereditary Partial Effective Functionals 487
4.1 The problem 488
4.2 The method 491
5 Invertible terms in λ-calculus 492
5.1 The problem 493
5.2 The method 494

Géométrie et syntaxe : les ensembles semi-algébriques et leurs généralisations ··· *M. F. Roy* **503**
1 Sturm-Sylvester 503
2 Les ensembles semi-algébriques 504
3 o-minimalité 506

A propos de l'imbrication des concepts analytiques et géométriques : le livre de Vitushkin sur la complexité du problème des tables ··· *M. Zerner* **511**
1 Introduction 511
2 Le problème des tables et sa transformation 512
3 La variation d'un ensemble et les inégalités fondamentales 514
4 Le problème du passage à la dimension infinie 517
5 En guise de conclusion 518
6 Annexe 1 : liste des chapitres 520

Geometry and Physics of the 20^{th} Century

Michael Atiyah

1 Historical and Philosophical Remarks

My aim in this lecture is not to survey in detail the history of geometry and physics in the 20^{th} century but to make some general remarks about the nature of geometry and its relation to physics, as exemplified by the past century.

The relation of mathematics to the reality of the external world, on the one hand, and to the workings of the human mind on the other have fascinated philosophers for centuries. Is mathematics simply derived from practical experience or is it a free creation of the human mind ? Why has mathematics been so successful as a tool for studying the real world, particularly in the field of physics ? Is God a mathematician ?

My own view is that such questions cannot be answered without a detailed examination of some of the ways in which mathematics and physics interact.

A historical perspective, following the evolution of ideas, is the most fruitful since it does not presuppose that our notions are frozen in time and have permanent validity.

While mathematics and physics have been progressing by leaps and bounds, our knowledge of the third part of the triangle, the workings of the human mind, is much more limited and still in its infancy. As we learn more about the neuro-physiology of thought our ideas will change, and the very questions that I posed at the beginning will seem irrelevant or at least improperly posed. At the present time our understanding of the way the brain works is beginning to make real progress and I expect that, within a few decades, this will shed much light on the age-old philosophical questions. But, for the time being, we are stuck in our ignorance and I will have to make my remarks in the traditional framework.

2 Space and Time

Let me begin with a simplified diagram that relates mathematics to physics.

$$\text{Mathematics} \left\{ \begin{array}{ll} \text{Geometry} & \text{Space} \\ \text{Algebra} & \text{Time} \\ \text{Analysis} & \text{Continuum} \end{array} \right\} \text{Physics}$$

The three great branches of mathematics are, in historical order, Geometry, Algebra and Analysis. Geometry we owe essentially to Greek civilization, Algebra is of Indo-Arab origin and Analysis (or the Calculus) was the creation of Newton and Leibniz, ushering in the modern era. Physics is the study of the space-time continuum, in all its aspects, and these three components can be reasonably paired off with their mathematical counterparts as illustrated.

There is little argument about Geometry being the study of space, but it is perhaps less obvious that Algebra is the study of time. But, any algebraic system involves the performance of sequential operations (addition, multiplication, etc.) and these are conceived as being performed one after the other. In other words Algebra requires time for its meaning (even if we usually only need discrete instants of time). The spatial dimensions are irrelevant. By contrast Geometry can be perfectly well conceived of in a static universe. For all practical human purposes the velocity of light is infinite so that pictures are seen instantaneously.

Of course in mathematics, Algebra and Geometry are clearly linked, while in physics space-time is (ever since Einstein) viewed as a unity. Analysis is required to describe dynamics, where spatial configurations evolve in time, and the true nature of the continuum and its significance both in mathematics and in physics has been one of the deepest questions faced.

3 Probing the Unknown

Let me now compare the different ways in which mathematics and physics progress, how we go from the known into the unknown.

Mathematicians indulge their imagination and fancy in moving forwards. Consider, for example, the introduction of such notions as $\sqrt{-1}$, higher dimensions or ∞ . To a physicist, with his feet on the ground, all this is "absolute nonsense", in clear contradiction with reality. Certainly, at the beginning, mathematicians themselves were very uncertain about the meaning of such non-existent entities and apologetic about their use. However they were persuaded eventually by the utility of these ideas. "Imaginary" numbers proved their worth, even for problems in the real domain. Similarly higher-dimensions (which Sylvester wanted to call "inconceivable") could be used to shed light on three dimensions, while a good understanding of infinity was essential to the calculus with all its ramifications.

So mathematical imagination, going outside the obvious physical world has paid off and even physicists have now reaped the benefit of this freedom.

In parallel, physicists frequently speculate about the unknown and introduce notions which are mathematically meaningless, such as the notorious Feynman integral. Again it is the utility of these notions in real physics that justifies their use, in the experience of physicists. By contract mathematicians view such speculation as suspicious, but history has shown that too rigid an attitude can hold up progress.

All this simply illustrates the different world views of mathematicians and physicists. To a mathematician reality is grounded in logic, and speculative ideas need have no relation with the external world so long as they are logically consistent. To a physicist, logic is secondary and experience is primary. Speculative ideas may have logical holes but as long as they are grounded in experience they are permitted.

In the dialogue between mathematicians and physicists, especially at the present time, the ground rules need to be well understood. Fruitful interaction will only emerge from mutual tolerance and an appreciation of the other viewpoint.

4 Scales

As mankind has delved further into the mysteries of the natural world we have learnt that human beings live at an intermediate scale. There are vastly larger scales, by many powers of 10, in the galaxies and many smaller scales, by many negative powers of 10, in the microscopic and atomic constituents of matter. One we study with telescopes, the other with microscopes and both of these basic instruments now come in many forms from radio telescopes to electronic microscopes.

We are now exploring the natural world, at both extremities, with tools of increasing refinement, and our understanding of Geometry gets modified as we progress.

In earlier times, rigid measuring rods were the standing tool for exploring space and Euclidean geometry emerged. Later we learnt to use light signals, linear accelerators and we also learnt about quantum effects. Our understanding of the nature of physical space altered and this has had an effect on geometrical thinking. This will become clearer when we move on to consider specific examples

5 Mathematics ahead of Physics

I want now to list a number of areas where mathematical development, inspired by creative imagination, has been well ahead of its physical applications which have often come decades later.

(i) **Curved Space**. Riemann's conception of space as being possibly curved was triumphantky indicated when it was taken up by Einstein in his theory of General Relativity.

(ii) **Lie Groups.** The theory developed by Lie and further extended by Elie Cartan has provided the framework of modern elementary particle physics with its idea of "internal symmetries".

(iii) **Higher Dimensions.** In the 19th century geometers explored the mysteries of higher dimensions, the "inconceivable" of Sylvester. Kaluza-Klein theory put forward in the 1930's is now fashionable as an essential framework for string theory.

(iv) **Fibre Bundles.** The mathematical treatment of fibre bundles was one of the most important aspects of differential geometry and topology in the 20th century. Having its roots in the work of Elie Cartan it received a major boost from Hermann Weyl's introduction of the notion of gauge theory in physics. This has now blossomed into a key part of modern physics. The honours in this case are really shared between geometers and physicists, with Weyl as the linchpin.

(v) **Spinors.** These have their roots in the work of Hamilton and Clifford, subsequently developed by Elie Cartan and Hermann Weyl . Their introduction into physics by Dirac has proved absolutely basic.

(vi) **Exterior Algebras.** Introduced by Grassmann in geometry and then applied by de Rham to differential forms, these provide the essential framework for Fermions.

(vii) **Non-Commutative Algebra.** This was the great breakthrough in 19th century algebra, pioneered by Hamilton and Cayley with the introduction of matrices. It provided the natural framework for the Heisenberg uncertainty principle of quantum mechanics.

(viii) **Hilbert Space.** The Spectral theory of self-adjoint operators provided the mathematical basis of quantum mechanics.

(ix) **Special Holonomy.** Differential geometers had classified the possible holonomy groups of Riemannian manifolds and these are now being used by physicists in M-theory, especially the Calabi-Yau manifolds (with $SU(n)$-holonomy) and the 7-dimensional manifolds with G_2-holonomy.

6 Simultaneous Discovery

In addition to the case of gauge theory mentioned above there are several other cases where new ideas developed simultaneously, either in partnership or independently.

(i) **Solitons.** The story of the development of the theory of solitons is a long one and it involved physics, mathematics and computer calculations. Scott Russell in the 19th century saw the first recorded soliton in a canal in Scotland. Much later computer calculations for the KdV equation exhibited soliton solutions. After the work of Lax the algebraic formalism of integrable systems developed rapidly. At all times there was equal interest and activity by mathematicians, interested in the theory, and physicists (or engineers) interested in the applications. The original ideas in 2 dimensions (1-dimensional space plus time) were then extended to 3 and 4 dimensions, linking up with ideas in particles physics that were originally proposed by Skyrme.

(ii) **Anomalies.** This is a case where, for a decade or more, mathematicians and physicists were tackling essentially the same problem quite independently. My own work with Singer on the index of elliptic operators turned out to be closely related to the

anomalies of quantum field theory. These anomalies arise when a classical symmetry is destroyed at the quantum level. For chiral symmetries this corresponds to index problems where the index is sensitive to orientation (often just changing sign). Once the link was made interaction between geometers and field-theorists grew apace. There is nothing like a concrete problem at the heart of your subject to stir interest. Certainly this bridge from the index theorem to anomalies was a turning point in the relation of mathematics to physics - it has led on to an extended and mutually fruitful collaboration.

7 Physics ahead of Mathematics

I will now turn to the cases where physics has been (and still is) ahead of the mathematics. Of course this has to be understood in the following sense. Physical motivation leads to speculative theories which have little in the way of a rigorous foundation. Nevertheless the ideas, techniques and "results" can show mathematicians the way forward. It may take a long time to prove these "results", either by making rigorous the physical ideas, or more frequently, by finding an easier alternative path. Knowing the goal is half the problem - finding the route is often a matter of technique.

(i) **∞-dimensional representations.** The mathematical theory of group representations was fully developed first for finite groups and then for compact groups. It was physicists (Wigner and Bargmann) who first looked at Hilbert space representations of non-compact groups (such as the Lorentz group). After this the mathematicians (Gelfand, Harish-Chandra, Langlands) took it up with a vengeance and made it a central topic in Mathematics.

More recently physicists moved on to consider infinite-dimensional groups and their representations. Prime examples are loop groups and the diffeomorphism group of the circle. They arise naturally in 2-dimensional conformal field theory. These topics have since been taken up by many mathematicians and they have found striking applications to such things as the Macdonald identities, the Monster group and moduli spaces associated to Riemann surfaces.

(ii) **Maxwell Theory.** Maxwell's equations provided one of the main ideas behind Hodge's theory of harmonic forms. Some of the applications of Hodge theory in higher dimension are, interestingly enough, now being found in Kaluza-Klein geometry[1].

(iii) **Dirac Theory.** The introduction of the Dirac operator for the quantum study of the electron was a high point of the early quantum physics. But it was not until 40 years later that it was used extensively in differential geometry, where it was fundamental in the Atiyah-Singer index theory[2]. This in turn is linked up to anomalies as just explained above.

[1] (see §5, number (iii))
[2] see the lecture by Bourguignon.

(iv) **Super-symmetry**. The need to treat Bosons and Fermions in parallel has led physicists to an extensive technology of "Super-symmetry". As yet this has not been fully absorbed by mathematicians, though many aspects are familiar : $\mathbb{Z}_2$-graded algebras, exterior algebras, equivariant cohomology. Perhaps the next generation of mathematicians will integrate super-symmetry more systematically into the mainstream of mathematics. It should have a significant role in homological algebra and its close links with Clifford algebras are clearly important.

(v) **Quantum Cohomology.** Classical algebraic geometry was often involved with counting the number of common solutions of a set of polynomial equations. This was "counting points". Algebraic topology and intersection theory (or the cohomology ring) was developed by Lefschetz partly to provide a rigorous framework for "counting points" (e.g. justifying the Schubert calculus). Ideas from quantum theory have now led to procedures for "counting curves" and this has been christened "quantum cohomology". It has now been given a rigorous foundation in the framework of modern symplectic geometry, but the most spectacular results were already predicted (correctly) by physicists.

(vi) **Conformal Field Theory.** I already alluded to this earlier but I mention it again because of its close connection with (5) above. In particular it has led to the notion of mirror symmetry for Calabi-Yau threefolds. This is a beautiful and still uncompleted subject which includes "counting curves" as part of the story.

(vii) **QFT in 3-dimensions.** Vaughan Jones introduced his famous polynomial invariants, starting with the theory of von Neumann algebras. It was Witten who reinterpreted the Jones polynomials in terms of a 3-dimensional quantum-field theory based on the Chern-Simons action. This transformed the situation and, in particular, gave generalizations to all 3-manifolds. Much of Witten's approach has now been made rigorous by algebraic and combinatorial techniques, though the final word has yet to be said. In particular the relation with the work of Thurston has barely begun.

(viii) **QFT in 4-dimensions.** The Yang-Mills equations arising from non-Abelian gauge theories in physics were brilliantly exploited by Donaldson to produce subtle and powerful invariant of smooth 4-manifolds. Witten first interpreted them in terms of a QFT and then, with Seiberg, argued that this QFT was "dual" to another QFT based on an Abelian gauge field with spinors. Mathematicians then turned their attention to this new Seiberg-Witten equation with spectacular results.

All these examples clearly demonstrate that mathematics has been transformed in the past 25 years by ideas and techniques arising from quantum field theory. Put philosophically, we are finally learning the real nature of space. Probing it with quantum fields is revealing a hidden and beautiful structure.

In my opinion this impact of quantum theory on geometry is even larger than the input General Relativity had in the early decades of the 20$^{\text{th}}$ century. It goes deeper and further.

8 M-theory

Continuing the theme of the previous section I want now to talk about current ideas in physics, where they are leading, and what their mathematical impact will be.

The great aim of present research in physics is to provide a unified consistent framework that includes both General Relativity and Quantum Field Theory. For some time now (super-symmetric) string theory has provided the only plausible candidate, but it was known that there were actually five different string theories (in 10 dimensions). So it was a great breakthrough when it became apparent that these five theories all seemed to be equivalent, and were different perturbative expansions of a unique theory, now called M-theory, which lived in 11 dimensions and was also related to 11-dimensional super-gravity. The equivalences were non-perturbative and involved various "dualities" as for example the duality between Donaldson-Witten theory and Seiberg-Witten theory.

The dualities can be viewed as non-linear analogues of the Fourier Transform, and they occur only at the quantum level. Understanding these dualities properly remains an outstanding technical and conceptual problem.

In many ways, and I have argued this before, such dualities should be viewed as similar to the dualities required for un-abelian class-field theory (the Langlands programme). In both the non-linearity of the problem requires the passage to infinite-dimensions to provide the sought for duality.

Quite recently there have been interesting links found between M-theory and Connes' non-commutative geometry. This looks very promising. It is quite clear that the main challenge for the 21^{st} century will be to find a proper framework for M-theory and its dualities. The physical application is, of course, to provide the unified theory of Gravity and Matter. But the mathematical consequences will also be profound.

9 Relation of Mathematics to Physics ?

Let me end, as I began, with a few philosophical remarks. There are various ways to try to describe the relation of mathematics to physics. First we have what I will call the **Economics Picture**. Here we view mathematicians as workers who produce **models** for use in the physical world. According to economic theory, they can respond to and be influenced by **market forces,** i.e. by physicists. We mathematicians produce what is required and we alter our production in the light of changing needs.

The second description, and one which I prefer, is what might be called the **Floral Picture.** Here we view mathematicians as **gardeners** who breed new species of flowers and display them in gardens. Physicists are like the 19^{th} century **plant collectors** who explored the wilds to bring back exotic specimens to reinvigorate gardens.

I view myself as a gardener, and I am grateful to the plant-collectors.

On The Early History of the Periodicity Theorem

Raoul Bott

This paper will be concerned with the period 1949 to 1959 and will, on the whole, be an account of Marston Morses work of the 1920's, and how naturally its reinterpretation in the language of the 40's leads to the periodicity theorem of the homotopy groups of the Classical Stable Groups. I hope to convince you that from that perspective the periodicity theorem is simply an extension of the Freudenthal Suspension Theorem, albeit a quite miraculous one. My paper will therefore also be more personal than is possibly appropriate in this forum, but I find the attempt to play the historian on a larger than personal level much too daunting to undertake, even though, from previous experience, I am also well aware of the perils posed to personal reminiscence by the highly selective nature of our memories. Nevertheless, I hope that this account will be of some use. But please remember, that all I can really promise you here, is how the past appears to me now, some 50 years later.

From this distance then, mine is, to a large extent, an account of fortuitous circumstances : that is, of having been at the right place at the right time. And the first, and most decisive of these circumstances occurred in 1949 when it was my great good fortune to be invited to the Institute for Advanced Study as a freshly minted Post Doc. in applied mathematics from The Carnegie Institute of Technology.

My position at the Institute was highly unusual in that my previous work had dealt with matters in Electrical Engineering and my original assignment had been to write a book on that topic at the Institute. However, the atmosphere at the Institute proved much too stimulating for such a mundane task and I soon found myself, instead, exploring the many fascinating aspects of pure mathematics that were going on around me. Professor Marston Morse was in charge of the Post Docs that year and eventually, and with considerable trepidation, my conscience led me to confess this new preoccupation to him. But Morse immediately dispelled my apprehension. "At the Institute you do what you want to do!" was his characteristically forceful rejoinder to my confession, and so, absolved, I was now free to follow my bent both at the Institute and at the University. Still, throughout my two subsequent years in Princeton I had the slightly guilty feeling of having insinuated myself into this august body under false pretenses.

Actually the program of ′49 in Princeton turned out to be ideally suited for some one in my predicament. On the one hand Norman Steenrod was giving a course on Fiberbundles at the University, which though advanced in character, was nevertheless accessible to a newcomer in Topology, thanks to Steenrod's wonderful habit of starting from scratch whenever possible. On the other hand, Herman Weyl was conducting a seminar on Hodge theory at the Institute, for which my training in electrical networks proved to be some sort of finite dimensional analogue. Nevertheless it was hard going for me, and I could not have absorbed much of the material if I had not had my good friend Ernst Specker to fall back upon as a personal tutor. But in addition to these two courses I also soon found myself very much taken with the book that Marston Morse had given me during one of our quite frequent encounters. His splendid office was just around the corner from my tiny one and if we met on the stairs he would often invite me into his office for a short chat. His primary interest at that time was in Analysis, and in particular with properties of functions of bounded variation, a topic which did not appeal to me very much. But for Morse this work was a technical expedition in his single minded effort to extend the "Theory of Critical Points" to more and more general situations.

In any case, on one of these occasions he gave me a copy of his book on the "Calculus of Variations in the Large" and it was browsing in this volume that soon became my private preoccupation. I say private, because nobody else at the Institute or the University seemed to be interested in this subject at the time. Nobody, except Morse of course. Indeed it was hard to have a conversation with him that did not eventually end in an exhortation about the general importance of the "Theory of Critical Points". For he felt that his vision of Topology and Mathematics as a whole, was being abandoned in the contemporary climate. And he was particularly antagonistic to the sort of Topology that was being taught at the University by Lefschetz and Steenrod. "For ever the foundation, and never the Cathedral" he complained in his Kenion Lecture. And indeed it was an era when we youngsters all felt the allure of doing Mathematics by "Definition" rather than by "Hard Work". Samy Eilenberg was possibly the most vocal exponent of this new "Conceptual" way of doing Mathematics at that time, and I have always loved his dictum : "Doing Mathematics should be like floating down the river in a canoe". Or again : "If you have to dip your oar into the water, you have already done something wrong". In short, there was a great dichotomy in the attitude I encountered at the Institute and the University in this regard, and I quite enjoyed seeing the sparks fly. For instance Carl Ludwig Siegel, who was maybe the most conservative of my heroes at the Institute would never use the word Homomorphism without prefacing it with the slightly pejorative : "The so called". On the whole then it was my impression that with the possible exception of Herman Weyl the Institute had the "Hard Work" attitude, while the point of view at the University was much more "Functorised" and conceptually oriented.

For me the great attraction of Morse's book, written in the 1920's, was twofold. First of all I was attracted by the grandeur of his central theorem to the effect that any well posed Lagrangian problem on the space of paths on a positive dimensional sphere must have an infinite number of extrema. And secondly I was much taken by the conceptional simplicity of the proof. The lesson was clear :

> Topological complexity of a space forces critical points of functions defined on that space and furthermore, in fortuitous circumstances, one can determine the

homology of a space purely by a careful analysis of the critical behavior of a function on that space.

So, for instance, in the early part of his book, which was devoted to what we now call the Baby Morse Theory, I learned the following method for computing the homology of a real Projective space. Consider a generic nondegenerate quadratic function (Ax, x) on Euclidean $(n+1)$ space. Restricted to the unit sphere it then descends to give a function on the underlying projective space, and it is easy to see that the critical points of the function so considered is given by the eigen directions of the associated eigen value problem : $Ax = lx$. Generically there will be $(n+1)$ of these, and if the corresponding eigenvalues are arranged in increasing order :

$$\{l(1), \ldots l(n)\}$$

then at the k'th critical point, p, our function restricted to the subspace generated by the the eigen directions of the first k eigenvalues will clearly have a nondegenerate maximum at p.

Now from this fact alone Morse deduces that the mod 2 cohomology of projective n-space must have precisely one generator in each dimension up to dimension n. Note that all the concepts involved in these criteria were well in the range of my mathematical training as an applied mathematician, whereas the conclusion was well outside that range. And so were, for the most part, the methods and arguments in Morse's book. First of all Topology was new to me, and as Morse's book was written long before the Axioms of Eilenberg Steenrod and exact sequences as well as the other modern concepts that I was learning in Steenrod's course, I found even the underlying concepts of his text strange and hard to grasp. So the natural course for me to follow was to read the arguments in Morse's book and then try to understand them in terms of the Steenrod's language.

Now, one of the modern concepts he taught us was, of course, the construction of a new space from an old one by attaching a cell along its boundary sphere via a map of that sphere into the space in question, and at some point during that first year it became apparent to me that the homotopy theoretic essence of the Baby Morse Theory could be distilled into the following two assertions. Suppose, then, that f is a smooth function on the compact manifold M, and let $M(c)$ denote the subspace of M on which f has a value less than or equal to c. Then the following is true :

Theorem A. *Suppose that f has no critical values in the range $[a, b]$, with $a < b$. Then $M(b)$ and $M(a)$ are Diffeomorphic.*

Theorem B. *Suppose that f has a single nondegenerate critical point in the interior of the range $[a, b]$. Then $M(b)$ has the homotopy type of $M(a)$ with a cell attached.*

Let me recall here that a critical point of f is a point, p, where the differential of f vanishes, and that such a point is nondegenerate if and only if the second order jet of f at p is nondegenerate. Furthermore at such a nondegenerate critical point, p, one also has the following

Morse Lemma *In the vicinity of a nondegenerate critical point p of the function f there exists a coordinate system relative to which, f takes the form :*

$$f = f(p) + x_1^2 + \cdots x_k^2 - (y_1^2 + \cdots y_l^2).$$

In short, near a nondegenerate critical point $f - f(p)$ is a nondegenerate quadratic function. Up to linear equivalence this function is therefore completely determined by the number of its negative eigen values, and Morse called this number the Index of p as a critical point of f. With this convention understood, Theorem B is refined by the following assertion :

> Under the conditions of theorem B the dimension of the cell attached to $M(a)$ to obtain $M(b)$ up to homotopy type, is equal to the index of p as a critical point of f.

In symbols, one writes :

$$M(b) = M(a) \cup e^l \qquad\qquad l = \text{index of } f \text{ at } p.$$

In Morse's book the consequences of his deformations were usually phrased in terms of his famous inequalities relating the Betty numbers of M to the number of critical points of a "nondegenerate" function on M, that is, one that exhibited only nondegenerate critical points on all of M. But to my generation the above formulation seemed the better way of summing up the consequences of Morse's insight. On the one hand this formulation is easy to remember, and the Morse inequalities follow by standard exact sequence arguments. Furthermore standard arguments now also yield Morses "lacunary" principle :

> If no two critical point indexes of a nondegenerate function f on M differ by one, then the integral cohomology of M is a free group with one generator for each critical point p of f ; this generator to be assigned the dimension equal to the index of f at p.

Similarly one deduces his "completing" principle :

> Suppose that to each critical point p of the nondegenerate function f we can assign a closed submanifold $N(p)$ of M such that f restricted to $N(p)$ attains its absolute maximum at p. Then the fundamental classes of the $N(p)$ form a basis for the Mod 2 cohomology of M. Further more, if all these $N(p)$ are orientable, then their fundamental classes are a basis for the integral cohomology of M.

Note here that it was this principle that yielded the cohomology of projective space alluded to earlier. But more importantly, this formulation applies easily and naturally to Homotopy Theory as well. Indeed a quite elementary general position argument analogous to the one that proves that the homotopy groups of an n sphere vanish below dimension n, shows that :

> Attaching an n-cell to a space does not change its homotopy groups in dimension less than $n - 1$.

In the early part of Morse's book I also encountered the symmetric square of a space for the first time. It was not abstractly introduced, but rather appeared naturally in the context

of the generalized billiard ball problem. Here one deals with a smooth, closed convex region in $(n+1)$-space (the generalized billiard table) and seeks the number of closed trajectories with k sides that a billiard ball might traverse. This problem can be reduced to one of counting the number of critical points on appropriate spaces. For instance to treat the simplest case, when $k = 2$, observe that any two points, p and q on the boundary of our region determine a unique straight line $[p, q]$, joining them in the region, and going back and forth along this straight line would be a possible billiard trajectory if and only if the straight line in question meets the boundary at right angles. But this condition is seen to be equivalent to the condition that the length of the line segment $[p, q]$ be extremal. Because the boundary of our region is a sphere of dimension n, we may interpret this length function as a function on the product of two n-spheres, which is symmetric under the involution interchanging the two factors, and hence descends to the symmetric product of the n-sphere. Now the topology of a product of two spheres forces only four critical points on a nondegenerate function. However, as we will see in a moment, the topology of the symmetric product of the n-sphere forces $(n+1)$ critical points on any such function. Hence the number of expected trajectories of length 2 is $(n + 1)$. QED

Morses procedure for solving this problem is analogous to the one he used for projective space. He first considers the problem when the convex region in question is a general Ellipsoid in Euclidean $(n + 1)$ space. Its principal axes are then clearly the only possible billiard trajectories with precisely two sides, and there are clearly $(n + 1)$ of them. Furthermore an easy explicit computation shows that the critical points of the length function on the Symmetric Square of n-sphere which correspond to these principal axes are all nondegenerate. But note now that clearly none of them constitute a minimum of our length function ! And indeed a short second order computation yields the indexes of these principal axes to be the set :

$$\{2n, 2n - 1, \ldots n\}.$$

To apply the completing principle argument here, Morse therefore had to work harder than he had to in the case of the projective space. But he managed, and so computed, what in contemporary language is the mod 2 cohomology groups of the symmetric square of a sphere modulo its diagonal.

By the way, these spaces were of considerable interest in their own at that time. On the one hand they were examples of the Smith theory dealing with fixed points of involutions and on the other, they were clearly somehow related to the newly discovered Steenrod Squares. In fact, my first serious paper in pure mathematics dealt precisely with this aspect of the symmetric squares of spaces.

In any case this was a prime example of how the Morse theory provided me with a private outlook on what was going on around me. And it was also the occasion for my first modest contribution to Critical Point Theory. That is, the introduction of the concept of a nondegenerate critical manifold. The definition is as follows :

> N is a nondegenerate critical manifold of the smooth function, f, if and only if N is a closed submanifold of M on which the differential of f vanishes and its second order jet has nullity equal to the dimension of N.

At such critical sets, the second order jet of f induces a nondegenerate quadratic form

on the normal bundle $v(N)$, of N in M, and once we chose a Riemann structure for $v(N)$, this form naturally translates into a non singular selfadjoint automorphism $H(f)$, of $v(N)$.

This "Hessian of f" therefore splits $v(N)$ into the direct sum of the two bundles generated by the eigen directions of $H(f)$ with positive and negative eigen values, respectively. Thus, along a nondegenerate critical manifold we have the decompostion :

$$v(N) = v(N)^+ + v(N)^-,$$

and the fiber dimension of $v(N)^-$ now serves to extend the notion of "index" to this more general situation. We call $v(N)^-$ the negative bundle of f along N. Note also that this index of N is constant on each component of N.

The impetus for this piece of "Mathematics by Definition" came about when, in the instance at hand, I asked myself what would happen if one replaced the generic Ellipsoid with the round Sphere. It is then quite obvious that the corresponding "Square of the length of the Chord" function on the symmetric square, would have only two critical values : A minimum on the diagonal and a maximum along the projective space $P(n)$, corresponding to the chords whose end points were antipodes of each other. It follows that the 2nd order jet of this function along its critical sets had to have a null space of dimension at least equal to the dimension of these submanifolds. But a short computation then convinced me that in this example the nullity was precisely equal to that dimension on both the critical sets. In short here we had an example of a function which was "nondegenerate in the extended sense". That is, all of its critical sets were nondegenerate critical submanifolds.

And now the Bundle Theory I was learning in Steenrod's course soon suggested the following appropriate extension of Theorem B :

Theorem C. *Suppose that f has a single connected nondegenerate critical manifold N in the interior of the range $[a, b]$. Then the homotopy type of $M(b)$ is obtained by attaching the unit disc bundle of the "negative" bundle $v(N)^-$ to $M(a)$.*

In symbols we write :

$$M(b) = M(a) \cup e\{v(N)^-\}$$

The cohomological consequences of Theorem C fitted very well with the early *Compte-Rendus Notes* of Thom on the cohomology of sphere bundles and in particular with the Thom isomorphism, according to which the mod 2 cohomology of a disc bundle over a space — modulo its boundary — was isomorphic to the cohomology of the base, but shifted by the dimension of the fiber. For instance, whereas in the Morse Polynomial of a nondegenerate function each critical point contributes t^l where l is the index of the point, Theorem C and the Thom isomorphism now imply that for functions nondegenerate in the extended sense each critical manifold N contributes t^l times the $P(t^l, N)$ with l the index of N and $P(t, N)$ the Poincaré polynomial of N.

In homotopy theory, Theorem C has the following as an easy consequence :

> Under the conditions of Theorem C the homotopy groups of $M(b)$ are equal to the homotopy groups of $M(a)$ in dimensions less than the index of N minus 1.

Indeed this can be seen by perturbing the function f in a vicinity of N, via a nondegenerate function g on N to a nondegenerate function f' on M whose critical points on M are

precisely the critical points of g on N, and whose indexes are the sum of their index qua critical points of g on N, and the index of N qua critical manifold of f on M. Consequently, by Theorem B, $M(b)$ is obtained from $M(a)$ by successively attaching cells all of dimension greater or equal to $l(N)$. QED.

This piece of "Mathematics by Definition" is hardly in a class with the Serre's definition of a fiber-space or Grothendiek's definition of K-theory, but it has nevertheless been of some use, and my special consciousness of it has served me well over the years. In particular, it eventually led me to see the Periodicity Theorem for the Unitary group as something that was completely obvious and had been staring me in the face for several years! But back in $'49$–$'50$, Theorem C was mainly my private piece of information. When I mentioned it to Morse he was complimentary but I could sense that he considered it a trivial extension of his ideas and insights. And so it is, really. For in Morse's book we find all manner of general position arguments, including what we now call Sard's Theorem. So the homotopy consequence of a nondegenerate critical manifold would have been quite natural to him, and I believe he once showed me a lemma in his book which essentially implied it. For Morse had, of course, proved that the Nondegenerate functions were generic, so that if one blows on any function it becomes a "Morse function", as we now commonly call nondegenerate ones. So the salient property to remember about the "Morse-Bott" functions, as one now often refers to functions that are nondegenerate in the extended sense, is that when blown upon appropriately they produce a Morse function where each critical manifold is replaced by critical points whose indexes equal or exceed its own.

I have spent so much time on these elementary aspects of Morse theory because once these are grasped it is very easy to formulate those consequences of the "Morse Theory Proper" that are pertinent to Homotopy Theory. For in that context we can dispense with a general Langrangian discussion, and limit ourselves to the consequences of Morse's work in the Riemannian case. The question to be addressed is then simply the following one :

> To what extent do the ideas and concepts that we have already discussed apply to the Energy Function on the Loop Space of a manifold ?

More precisely, consider the space of, say, piecewise smooth maps of the unit interval I, into the complete Riemannian manifold M. If we denote this space, in an appropriate topology to be discussed later, by Ω, and refer to its points as paths, then the projection of of a path onto its end points defines a projection of Ω to $M \times M$, and the fiber of this projection over (p, q) will be denoted by $\Omega(p, q)$. We assume M to be connected. Then it is an elementary fact of Homotopy theory that :

> The Homotopy groups of $\Omega(p, q)$ are those of M, but shifted down by one dimension.

In short all these spaces, as p and q range over M, are of the same weak homotopy type, and so all have the same homotopy groups, namely those of M shifted down by one. Thus the fundamental group of M corresponds to the components of $\Omega(p, q)$ and so on. The underlying Riemann Structure of M now defines an Energy function on Ω, which we will denote by S, as it plays the role of the "Action" in general Lagrangian theory. S assigns to a path $u(t)$ the integral :

$$S = 1/2 \int_I \langle u', u' \rangle \, dt$$

with the prime denoting the derivative of u relative to t.

The critical points of this function, restricted to $\Omega(p,q)$, now are well known to correspond precisely to the geodesics joining p to q in M, and their study therefore brings us to the very heart of Riemannian theory, so that all the definitions of Morse applying to Calculus of Variation problems of this type now acquire a beautiful geometric significance.

In this essentially infinite dimensional situation, the Tangent Space at a path u, should, of course, be taken to be the space of piecewise smooth vector fields along u which vanish at the end points of u, and the first step of the Morse Theory Proper is to argue that the role of the Hessian of S at a geodesic u, is replaced by a self-adjoint differential operator in this tangent space. And indeed, the classical formulas of second variation show that in this Riemannian situation, the appropriate Hessian is given by the differential operator :

$$Y \to -DDY + R(X,Y)X,$$

associated to the Jacoby Differential Equation along u. Here Y denotes a tangent field along the extremal u, while D stands for covariant differentiation in the tangent direction, X, of u, and R is the Riemannian curvature of M.

The Sturm theory now teaches one that although the spectrum of this Hessian is unbounded, it is bounded from below, and discrete with finite multiplicity. Thus the index of S at u is well defined as the number of negative eigenvalues of $H(S)$ and its nullity as the multiplicity of zero as an eigenvalue of $H(S)$.

Correspondingly one can define nondegeneracy of a geodesic by the requirement that this nullity be zero, and nondegeneracy of a critical submanifold N of geodesics in $\Omega(p,q)$ by the requirement that the nullity of N be equal to its dimension. In short, all the concepts of the Baby Morse theory extend naturally to this situation and the Morse Theory Proper, may now be stated very simply as the assertion that Theorems A, B, and C remain valid for the function S in this infinite dimensional context. From the topologists point of view, the primary consequence of this theory is then that :

> A second order understanding of the geodessics joining p to q in M, can, in fortuitous circumstances, lead to an understanding of the loop space of M as a Cell complex.

For example, let us take M to be the two sphere in its usual metric inherited from 3-space. Chose p to be the north pole, and q to be a nearby point in the northern hemisphere. The geodesics in $\Omega(p,q)$ then consist of the subarcs on the great circle containing p, and joining p to q. To keep book on all of these, let p' be the south pole on this circle. It is then clear that the geodesic arcs in question are uniquely specified by the number of times they contain p'. Thus the shortest of them does not contain p' at all, the next longest comes around the other way and contains p' once and so on. If we next examine the Jacoby equations appropriate to this situation we find that all the geodesics in $\Omega(p,q)$ are nondegenerate, and that the index of the k'th one in our count is precisely k. By Theorems A and B we conclude that the loop space of the 2-sphere is obtained from a point by successively attaching cells, one in each dimension. In symbols we write :

$$\Omega S^2 = pt \cup e^1 \cup e^2 \cup \cdots .$$

For the higher dimensional spheres this pattern is seen to persist, except that now the index of the geodesic arc containing k copies of p' is equal to $k(n-1)$. Thus we arrive at the cellular picture :

$$\Omega S^n = pt \cup e^{(n-1)} \cup e^{2(n-1)} \cup \cdots .$$

Note the interesting fact that this decomposition, pure and simple, computes the integral cohomology of ΩS^n as having a free generator in each dimension divisible by $(n-1)$, provided n is greater than 2, as a consequence of the lacunary principle. The assertion actually remains valid for the 2-sphere also, but then needs further geometric insight and eventually follows by an application of the completion principle.

Grosso modo this describes Morse's procedure for computing the homology of the loop space of the spheres, and applying his theory, so to speak, the other way around, he deduces that in any metric on the n-Sphere with $n \geq 1$, there must be an infinite number of geodesics joining any two given points. This follows readily if we assume that all the geodesics in question are nondegenerate, for then the infinite number of cohomology classes force an infinite number of critical points. To obtain the general result he then augments this argument with appropriate general position arguments.

I expect that Morse's remarkable computation of the cohomology of these loop space attracted so little attention in the ′30's and ′40's because there seemed to be no direct consequence for topology or homotopy theory in that computation so far, and Morse's own emphasis was always the other way around. However, if we now pass on to the 'extended' form of nondegeneracy one immediately obtains the central result of the early homotopy theory, namely, the "Freudenthal Suspension Theorem."

Indeed, let us now consider the loop space on the n-sphere, but this time in its guise as the space of paths joining p to its antipode p'! The geodesics joining p to p' now clearly fall into manifolds which themselves are $(n-1)$ spheres. Furthermore an easy computation, to which I will return later, now establishes that they are all nondegenerate, and that their indexes, as such, run through the set of integers of the form $\{k(n-1)\}$. By Theorem C and the estimate of its corollary we therefore obtain the decomposition :

$$\Omega S^n = S^{n-1} \cup \text{ cells of dimension greater than or equal to } 2(n-1).$$

But then it follows that this loop space has homotopy groups equal to those of the $(n-1)$ sphere in dimensions less than $2n-3$. In view of the fact that taking the loop space lowers homotopy by one, one finds that the homotopy groups of S^n and S^{n-1} agree in low dimensions, that is in the range less than $2n-3$, in short, precisely the result of Freudenthal — dating back to 1937. I had heard about this theorem in Steenrod's course, and by the end of ′49 I think that this line of proof was my private way of understanding it. Later, this cell decomposition of the loop space of the spheres was obtained by Ioan James without Morse theoretical methods and still later John Milnor described a cell decomposition for the loop space of the suspension of any cell complex.

But in the Princeton of ′49–′50, I do not recall any mention of loop spaces, except in my discussions with Morse. The French Revolution started by Serre's thesis had not yet occurred. For instance, in Steenrod's course the higher homotopy of the Spheres were a great mystery. All we knew was that they vanished in low dimensions and that the nth homotopy group of the S^n was Z, and the $(n+1)$st one was $Z/2$. Concerning the higher ones, all we knew

was that they were countable Abelian groups. In fact I distinctly remember the day when Steenrod entered the classroom triumphantly to announce that the $(n+2)$nd homotopy group of S^n had just been determined by Pontryagin to be 0! — only to return a week later with George Whitehead's correct value of $Z/2$. So, at that time, common homotopy knowledge was to a large extent limited to the more or less standard consequences of the long homotopy exact sequence of a Fiber Bundle. But amongst these there was certainly the fact that the "low" homotopy groups of the Unitary group, $U(n)$, were independent of n, so that it made sense to speak of the "Stable Homotopy group of U". Indeed this result follows directly from the fact that $U(n)$ fibers over the $(2n-1)$ sphere with $U(n-1)$ as fiber. I was also very much aware that a similar stability result was true for the Grassmanian, $U(2n)/U(n)\times U(n)$, and that these stable homotopy groups of "$U/U\times U$" were equal to those of U *but shifted up by one*! For these spaces played a fundamental role in the classification theory of the Unitary bundles.

In short, by the end of my stay at Princeton in the summer of $'51$, most of the ingredients needed to prove the periodicity theorem for the Unitary group were at my disposal!

All, save two. First of all, there was absolutely no reason to conjecture such a result and secondly my understanding of the geometry of compact Lie Groups was nonexistent. And so it was the only in 1956–7, during my second two year stint at the Institute that all the necessary ingredients finally came together one fine afternoon in my, now somewhat larger, office there. All that was needed was for me to contemplate the loop space of $U(2n)$ in the guise of the space of paths on $U(2n)$ which connect the identity in $U(2n)$ to an element J which is the identity on an n dimensional subspace V, and minus the identity on its orthogonal complement. For then the set of shortest geodesics in this loop space of $U(2n)$ is easily seen to be the Grassmanian $U(2n)/U(n)\times U(n)$. Furthermore it turns out that in this instance the energy function is nondegenerate in the extended sense, and that it is relatively easy to compute all the indexes of the "higher" nondegenerate critical manifolds in question. The upshot was then the following cellular description of the weak homotopy Type of $\Omega U(2n)$:

$$\Omega U(2n) = U(2n)/U(n)\times U(n) \ \cup \ \text{cells of dimension greater than } 2n.$$

Because the space on the left has shifted the homotopy of $U(2n)$ *one dimension down*, while the one on the right has shifted the low homotopy groups of $U(2n)$ *up by one dimension*, the periodicity theorem follows from this formula by simply letting n tend to infinity. QED

The primary occasion for this computation was the controversy that had arisen at that time between Homotopy Theorists and the computations of Borel and Hirzebruch. By the mid-fifties the climate of Homotopy theory had of course completely changed. Serre's adaptation of Leray's Spectral sequence to Homotopy Theory in the early $'50$'s had introduced completely new and sophisticated algebraic techniques to the subject, and one now had, for the first time, the feeling that the low homotopy groups of reasonably simple spaces could be computed.

Serre's Thesis was a quite especial delight to me, because it provided such a new perspective on Loop Spaces and in fact made them the cornerstone for the computation of homotopy groups. Roughly, Serre's method can be summarized as follows :

> The Spectral sequence of a fibering gives one a handle for computing the homology of the loop space of X from the homology of X, and as this construction lowers the homotopy groups, a sufficiently large iteration of the loop construction brings the higher homotopy groups into a range where the Hurewitz theorem applies, so that homotopy and homology groups are identical.

In fact Serre's Thesis is one of the very few papers that I have ever studied in great detail. This occurred during my first year at Ann Arbor in 1951, largely thanks to the quiet but determined influence of Hans Samelson, my newly found friend and mentor there. We ran a truly detailed and workman like seminar on the subject, which soon cut into the attendance of the graduate students, but taught us the subject sufficiently to make some contributions of our own to it later on that year.

In any case, by 1956 these techniques had led the Homotopy Theorists to announce that the tenth stable homotopy group of the unitary group was $Z/3Z$. On the other hand Borel and Hirzebruch's work on Characteristic Classes had convinced them that this group should vanish! This controversy immediately attracted my attention, for I sensed that the Morse Theoretic methods that I and Samelson had been applying to the Loop Space of a compact Lie Group over the past few years should be able to throw quite independent light on the subject.

Eventually my new friend Arnold Shapiro, whose universal interest in Mathematics was legendary, and to whom I had described the Morse Theoretic methods Samelson and I had developed to study the ring structure of flag manifolds, found a way to an independent test of the controversy. It was a complicated procedure involving the exceptional group $G(2)$, and the computation of Steenrod operations at the prime 3, in the cohomology of its Flag manifold. I distinctly remember that Arnold and I set siege to the small lecture hall in the Math building during a hot humid weekend for this purpose, and at the end of our arduous labour we came out convinced that Borel and Hirzebruch were right.

Here I must pause to remind you of my warning at the start of this narrative, concerning the selective nature of our memories. For, a few years ago when Borel and I chanced to meet in Hong Kong, and we were discussing the good old days, he 'reminded" me that actually this computation of Arnold and mine had at first come out in favor of the Homotopy Theorists! And that it was he who pointed out the error in our argument to us when we ran it by him for confirmation! I am convinced that Borel is correct, but, alas, I have not the slightest memory of this preliminary misstep! However such a lacuna in ones memory is quite consistent with my, now often tested theory, that on the whole we tend to forget incidents in which we play less than heroic parts, unless, of course, they were embarrassing in the extreme.

In any case, the final outcome of this computation convinced me that the 10'th stable homotopy group of U vanished. But then the periodic pattern of the known homotopy of U was striking, and it occurred to me that this pattern might persist. And it seems to me now, that once this thought had entered my head, the computation described above, was nearly self evident. For by that time I had already spent several years mining the quite miraculous relation of the Morse Theory to the Topology of the loop space of a compact Lie Group.

The start of this realization occurred in 1953–4, and I am now unsure what brought it about. All I recall is that at some point Samelson suggested that I read the paper of Stiefel concerning the Diagram of a compact Lie Group. Ironically, Samelson went to the Institute at Princeton the very year my interest turned in this direction, otherwise he would have

certainly been my ideal personal instructor and collaborator. In any case, I found Stiefel's global account much more suited to my taste than the other more Lie theoretic ones I had attempted before, and was soon overtaken with wonder how so many of the secrets of a compact connected Lie Group G, could be read off from a maximal Torus T of G and the disposition of the sub tori in T corresponding to the kernels of its root-homomorphisms.

Recall here that these arise from the decomposition of the tangent space of G/T at the coset T, via the Adjoint Representation. Indeed, the natural action of T on this space decomposes it into a finite set of 2-planes, $E(a)$, $(a = 1, 2, \ldots, m)$, called the root spaces of G, on which T is represented by rotations, and in terms of these the Diagram of G now simply consists of T together with the set of sub tori $T(a)$ which act trivially on $E(a)$.

> This Diagram of G is then possibly best displayed as a configuration of planes in the vector space T', corresponding to the universal cover of T, in which the lattice corresponding to the identity in T is also distinguished.

For instance, in Figure 1 below one finds the Diagram of SU(2), that is, the group of unit Quaternions so displayed. In Figure 2 I have drawn the corresponding picture for the

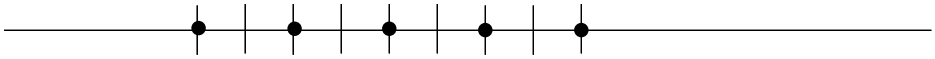


FIG. 1 – Diagram of SU(2)

exceptional group $G(2)$, and I must admit that to this day I am filled with awe that this diagram completely holds the secrets of this subtle exceptional group.

Contemplating these pictures and the others corresponding to the groups whose maximal Tori had dimension less than two, it soon became apparent to me that they were really much more naturally connected to the topology of the Loop space of G than to the topology of G itself! This comes about because of the great symmetry of the Riemannian geometry of a compact Lie Group in a left and right invariant metric, and its intimate connection to its purely group theoretic properties. Thus the geodesics in such a metric will always correspond to the left and right translates of the one parameter subgroups, and subgroups are geodesic submanifolds.

Now the geodesics on a Torus are of course easy to fathom : they appear as straight lines in the universal cover, and the first indication that the Morse theory is so appropriate in this situation is that :

> Let p be a point of T which does not lie on any of the root tori of the Diagram of G. Thenall the geodesics joining p to the identity in G lie on T.

Interpreted in our universal cover of T this statement therefore gives us a complete oversight of the geodesics in G joining p to the identity in G, as the set of straight lines joining the origin of T' to the lattice of points p' which correspond to p.

And next one can ascertain the following, even more satisfactory, state of affairs. Let p be any point in G, and consider the energy function, S, on the Loop Space of G in the guise of paths joining p to the identity in G. Then :

The Energy function, S, is nondegenerate in the extended sense.
The nondegenerate critical manifolds of S are in one-to-one correspondence with the straight line segments, $\{l\}$, joining the identity to the lattice of pre-images p' of p in T'.

And most important of all :

The critical manifold associated to the line segment, l, is diffeomorphic to the homogeneous space $G/$ Centraliser of l. Furthermore its index is equal to two times the number of intersections of the planes of the Diagram of G with the interior points of l.

The proof of all these assertions depends on Marston Morse's beautiful Index theorem, and, of course, the great symmetry of the geometry of a compact Lie group. But before we turn to that topic, note that these properties now immediately imply that the *loopspace of G is always free of Torsion* and one can read off the Betti numbers of ΩG by simply counting the number of times the straight lines joining the identity to the lattice points in the inverse image of p cross the planes of the diagram. Using the symmetry of the situation these lines can be brought into a one-to-one correspondence with the "Alcoves" of a "Fundamental Chamber" of the Diagram of G. In this way then the numbers I have indicated in the two figures completely specify the Betti number of the loopspaces of these respective groups.

Returning to the Index Theorem of Morse, recall that the definition of the index of a critical point in the Morse Theory Proper, was defined as the number of negative Eigenvalues of the Hessian of the function under consideration. The Morse Index Theorem now evaluates this number in terms of the behaviour of the vector fields along the extremal which are annihilated by the Hessian, and in our Riemannian context this amounts to the following assertion.

Recall first that if s is a geodesic segment on a Riemann manifold M, then the Jacoby equations along s are given by :

$$DDY - R(X, Y)X = 0,$$

where, as before, X describes the tangent field along s, D stands for the covariant derivative in the direction of X, and R for the curvature, relative to the Levicivita connection on M. Solutions of this equation are called Jacoby fields, and the end points of s are defined to be *conjugate* to each other along s if and only if there exists a non trivial Jacoby field which vanishes at *both* of these end-points. Further one then defines their "multiplicity" as the dimension of the space of Jacoby fields satisfying that boundary condition. With these concepts recalled, the Morse Index Theorem specializes to the following assertion in this Riemannian context :

A geodesic segment, s, corresponds to a nondegenerate critical point of the Energy function, if and only if its end points are not conjugate along s. Furthermore the index of such a nondegenerate critical point is given by the sum of the number of conjugate points of an end point of s which lie in the interior of the segment s, counted with their multiplicity.

The pertinence of this theorem of Morse to the matter at hand is surely clear. All that remained for me to do was to find the disposition of conjugate points along a geodesic in terms of the Diagram of the group. The appropriate concept for this translation turned out

to be one I called "Variational Completeness", and to describe it the following remarks will be useful.

Suppose that M is a Riemann manifold, and K is a group of isometries of M. Then every subset A of M has a "stabilizer", $K(A)$, consisting of the subgroup of K keeping A fixed pointwise. Also, if s is a non trivial geodesic segment we call s "transversal" if it is perpendicular to the orbits of K. It is an elementary but important fact that if s is perpendicular to the orbit of K at one pint, then it will be so at all of its points. So, for instance, a geodesic segment starting at a fixed point of the action is necessarily transversal. With this terminology understood, I called such an action "Variationally Complete", if and only if :

> Any Jacoby Field that vanishes at the end points of a transversal geodesic segment, s, is the restriction of an infinitesimal motion of K to s.

It is now a relatively easy matter to show that, on the one hand the *Adjoint Action* of a compact Lie Group on itself is *Variationaly Complete*, and on the other, that for such actions the disposition of conjugate points on a geodesic segment, s, starting at a fixed point, p, of the action is completely determined by the following dimension count :

> The multiplicity of conjugacy of the end points of such a geodesic segment s is equal to the dimension of the homogeneous space $K(q)/K(s)$, where q is the other end point of s.

Finally, it is also easy to see that we could have defined the Diagram of a group G by a similar dimension count. Namely, that :

> $D(G)$ consists precisely of those points p on the maximal Torus T of G, whose stabilisers under the adjoint action have dimension greater than the dimension of T.

These are then all the steps needed in my 1955 paper to make the translation from the index theorem of Morse to its group theoretic interpretation. Actually by the time of the publication of that paper, Samelson and I had already come to understand that for actions of this type, the Morse Theory for the Energy function on the space of paths on M, joining any point p on M to *any orbit of the action*, is always nondegenerate in the extended sense, and that in the nondegenerate case one has very natural "completing manifolds" for all the critical points, which we called their K-cycles.

But it is high time to return to the events of the spring of 1957, when in view of the foregoing, the periodicity Theorem for U became nearly self evident to me. In fact, it seemed so obvious that I was at first highly suspicious of it and kept checking it over and over. I do not think that I announced it to anyone until the summer of that year. Partly I felt insecure about the details of the Morse Theory Proper. There were all the troublesome questions that I have glossed over so far, like the appropriate topology to put on the space of paths, and how to summarize the many deformations involved in Morse's original account. The book of Seifert and Threlfall was my main guide there, but my final, somewhat different method took another year of work. But possibly, what really stopped me from announcing my news, was the feeling that I now was very close to a much grander view of periodicity. For, in complete analogy with the Unitary group the other classical groups $\mathrm{SO}(n)$ and $\mathrm{SP}(n)$ also had "stable" homotopy groups in the sense that their low dimensional homotopy groups

were independent of n and I was now hopeful that these stable values would also turn out to be periodic.

And eventually it did turn out that applying the Morse Theory along the same lines that had worked in the Unitary case, led straight to the 8-fold periodicity in these cases, and also yielded the 4 fold shift between them.

To start you on this road, consider the group $SO(2n)$. According to our method, we look for a point J, such that the set of minimal geodesics on $\mathrm{SO}(2n)$ which join the identity to J, have as large a dimension as possible. It then suggests itself to try a J which serves to define an almost complex structure on the underlying vector space. Consequently the set of minimal geodesics joining the identity to J is given by the homogenous space $\mathrm{SO}(2n)/U(n)$, and a routine check of the indexes of the longer critical manifolds now shows that the dimension of these tend to infinity with n.

Hence in the limit we can replace the loop space of $\mathrm{SO}(2n)$ by the homogenous space $\mathrm{SO}(2n)/U(n)$. In Symbols we write :

$$\Omega\mathrm{SO} = \mathrm{SO}/U.$$

At first sight it might appear that we have run into a dead end because $\mathrm{SO}(2n)/U(n)$ is not a Lie-Group. But note that it is manifestly a Symmetric space, and by that time Samelson and I had already come to understand that the pertinent theorems that made the Morse Theory so applicable to Lie Groups, actually also held in the larger category of Symmetric Spaces ! More precisely, we knew that the following assertions were valid :

If M is a Symmetric Space given in the form G/K, where K is the fixed pont set of an involution of the compact connected Lie Group G, then :

(1) The action of K on M is variationally complete.

(2) The energy function on the loop space of M is nondegenerate in the extended sense for any disposition of the end points.

(3) M contains a "maximal Torus T", together with a Diagram which plays a quite analogous role to the one it did in the Lie Group case.

To urge one on further, there is also the following marvelous fact :

(4) If M is a Symmetric Space and p and q are any two points on M, then the set of minimal geodesics joining p to q is again a symmetric space.

In view of this state of affairs it was clear how to proceed, and then a little searching soon provided the following symbolic extension of the formula above :

$$\begin{aligned} \Omega\{\mathrm{SO}/U\} &= U/\mathrm{SP} \\ \Omega\{U/\mathrm{SP}\} &= \mathrm{SP}/\mathrm{SP}\times\mathrm{SP}, \end{aligned}$$

from which we may deduce the 4 fold shift from SO to SP, and the corresponding formulae :

$$\begin{aligned} \Omega\mathrm{SP} &= \mathrm{SP}/U, \\ \Omega\mathrm{SP}/U &= U/O, \\ \Omega U/O &= O/O\times O, \end{aligned}$$

which yields the 4 fold shift the other way around.

I seem to remember that I became convinced of the 8 fold periodicity in these two cases during the early part of the summer of 1957, while we were moving to a new house in Ann Arbor, but via a different argument. The itterative one described above only became clear to me later and the details are described in my 1958 paper on the subject. However, I certainly have a very vivid memory of announcing the result at the University of Chicago later in the summer of ′57. It sticks so firmly in my mind, because, after my lecture one of the true experts in homotopy theory assured me that although he could not find a counter example at the moment, he was sure that the periodicity could not be true!

This really brings me to the end of my narrative. To carry on and outline the many other subsequent proofs of the periodicity, or its spectacular applications in its new guise as the corner stone of a generalized cohomology theory, due to Atiyah and Hirzebruch, or its subsequent intimate relationship to the Index Theorem of Atiyah and Singer, and all its K-Theoretic generalizations is quite beyond my scope. But I hope that these topics will someday be treated by more competent hands than mine. So rather, let me end here with just two remarks. The first is, that the whole Morse Theory Proper, and in particular this application to Topology fits much more naturally into the present friendly climate between Mathematics and Physics than it did 50 years ago. In fact, it now stands in a long line as a first, and rather elementary, "Physics Inspired" result in Topology. The second is, that, as this story amply illustrates, in the mysterious world of mathematical discovery ignorance can be a virtue, and — especially for us ordinary mortals — nothing beats being at the right place at the right time.

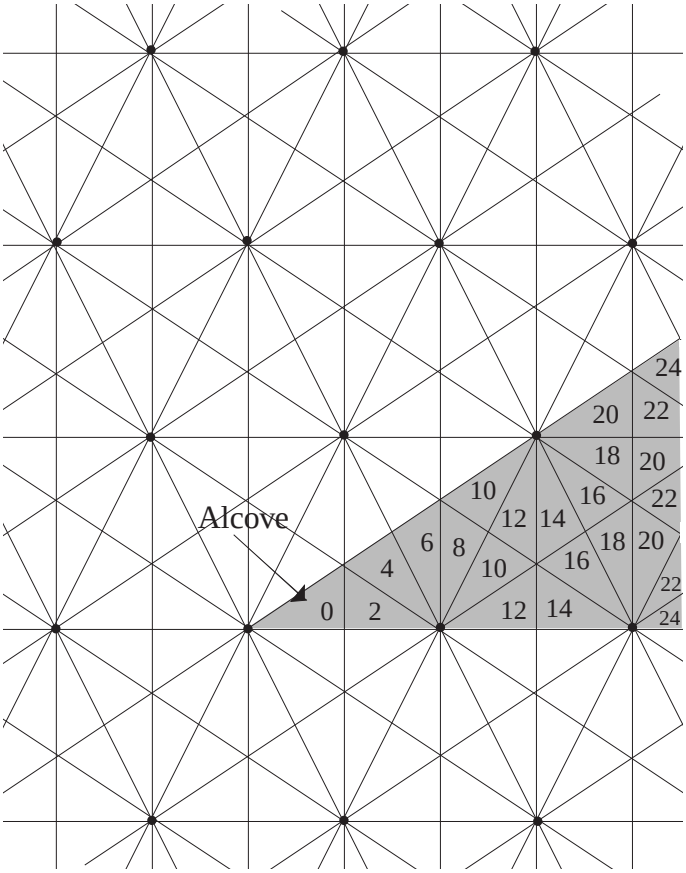


FIG. 2 – Diagram of $G(2)$

Bibliographie

[1] R. Bott, *An Application of the Morse Theory to the topology of Lie-Groups*, Bull. Soc. Math. France, **84** (1956), 255–281

[2] R. Bott and H. Samelson, *Applications of the Theory of Morse to Symmetric Spaces*, Amer. Jour. of Math., **80** (1958), 964–1024

[3] H. Freudenthal, *Uber die Klassen von Sphären abbildungen*, Comp. Math., **5** (1937), 299–314

[4] M. Morse, *Calculus of Variations in the Large*, Coll. Pub. Amer. Math. Soc. Coll. Publ., no. **18**, Providence, RI, 1934

[5] H. Seifert and W. Threlfall, *Variations Rechnung ih Grossen*, Teubner, 1938

[6] J. P. Serre, *Homologie Singulìere des Espaces Fibrés*, Ann. Math., **54** (1951), 425

[7] E. Stiefel, *"Ueber Eine Beziehung zwischen geschlossenen Lie'schen und discontinuierlichen Bewegsgruppeuklidischer Raüme und ihre Anwendung aufdie Aufzählung der einfachen Li'schen Gruppen*, Com. Math. Helv., **14** (1942), 350

VON ZAHLEN UND FIGUREN

Yu. I. Manin
Max–Planck–Institut für Mathematik, Bonn

1 Introduction

Geometry is a large subfield of mathematics, but also a label for a certain mindset of a practising mathematician. The same can be told about *Algebra* (understood here broadly, as the language of mathematics, as opposed to its content, and so including *Logic*.) A natural or acquired predilection towards geometric or algebraic thinking and respective mental objects is often expressed in strong pronouncements, like Hermann Weyl 's exhortation of "the devil of abstract algebra" allegedly struggling with "the angel of geometry" for the soul of each mathematical theory. (One is reminded of an even more sweeping truth : *"L'enfer – c'est les autres"*.)

Actually, the most fascinating thing about algebra and geometry is the way they struggle to help each other to emerge from the chaos of non–being, from those dark depths of subconscious where all roots of intellectual creativity reside. What one "sees" geometrically must be conveyed to others in words and symbols. If the resulting text can never be a perfect vehicle for the private and personal vision, the vision itself can never achieve maturity without being subject to the test of written speech. The latter is, after all, the basis of the social existence of mathematics.

A skillful use of the interpretative algebraic language possesses also a definite therapeutic quality. It allows one to fight the obsession which often accompanies contemplation of enigmatic Rohrschach blots of one's imagination.

When a significant new unit of meaning (technically, a mathematical definition or a mathematical fact) emerges from such a struggle, the mathematical community spends some

[0]Talk at the International Conference "Géométrie au vingtième siècle : 1930–2000", Paris, Institut Henri Poincaré, Sept. 2001. The title is a homage to Hans Rademacher and Otto Toeplitz whose book fascinated the author many years ago.

time elaborating all conceivable implications of this discovery. (As an example, imagine the development of the idea of a continuous function, or a Riemannian metric, or a structure sheaf.) Interiorized, these implications prepare new firm ground for further flights of imagination, and more often than not reveal the limitations of the initial formalization of the geometric intuition. Gradually the discrepancy between the limited scope of this unit of meaning and our newly educated and enhanced geometric vision becomes glaring, and the cycle repeats itself.

A very special role in this cyclic process is played by problems. The importance of a new theoretical development is generally judged by its success (or otherwise) in throwing light on an old problem or two. Conversely, a problem can stimulate the emergence of a new geometric vision expressing a sudden perception of a hidden analogy, as in André Weil's famous recognition of Lefschetz's formula in the context of algebraic equations over finite fields.

A problem/conjecture is usually represented by a short mathematical statement allowing a yes or no answer. "Yes" and "no" do not play symmetric roles : a positive answer to a good question usually validates a certain intuitive picture, whereas a "no" answer often shows only the limitations of this picture rather than its total lack of value. A "counterexample" disposing of a resistant conjecture can have a certain sportive value, but becomes really important only if pursued so far as to reveal some positive truth which has been escaping our understanding for some time.

The geometry of XXth century is a huge patchwork of ideas, visions, problems and its solutions. A brief list of platitudes I sketched above could be used to organize a narrative dedicated to the contemporary history of this discipline or its separate episodes. I have chosen instead to present at this conference a narrative based on my current work which at its key point remains purely conjectural, if not outright speculative.

I hope that the lack of a real mathematical breakthrough to report can be compensated in the context of this conference by certain freshness of perception accompanying such early stages of research. Besides, a discussion of partial successes and failures of this enterprise can serve as an illustration of some general issues of geometry and algebra.

Briefly, the research in question concerns explicit construction of numbers generating abelian extensions of algebraic number fields.

The archetypal problem is that of understanding abelian extensions of $\mathbb{Q}$. As we know after Kronecker and Weber, the maximal abelian extension of $\mathbb{Q}$ is generated by roots of unity. Roots of unity of degree n form in the complex plane vertices of a regular n–gone. Which of these n–gons can be constructed using only ruler and compass, was a famous problem solved by Gauss . His first publication dated April 18, 1796 (and *Tagebuch* entry of March 30, cf. [13]) is an announcement that a regular 17–gon has this property. Gauss was not quite 19 then ; apparently, this discovery prompted him to dedicate his life to mathematics.

One remarkable feature of Gauss ' result is the appearance of a hidden symmetry group. In fact, the definitions of a regular n–gon and ruler and compass constructions are given in terms of Euclidean plane geometry and make practically "evident" that the relevant symmetry group is that of rigid rotations $SO(2)$ (perhaps, extended by reflections and shifts). This conclusion turns out to be totally misleading : instead, one should rely upon

$Gal\,(\overline{\mathbb{Q}}/\mathbb{Q})$. The action of the latter group upon roots of unity of degree n factors through the maximal abelian quotient and is given by $\zeta \mapsto \zeta^k$, with k running over all $k\,mod\,n$ with $(k,n)=1$, whereas the action of the rotation group is given by $\zeta \mapsto \zeta_0\zeta$ with ζ_0 running over all n–th roots. Thus, $Gal\,(\overline{\mathbb{Q}}/\mathbb{Q})$ does not conserve angles between vertices which seem to be basic for the initial problem. Instead, it is compatible with addition and multiplication of complex numbers, and this property proves to be crucial.

This gem of classical mathematics contains in a nutshell some basic counterpoints of my presentation : geometry and spatial imagination vs refined language of algebra ; physics (kinematics of solid bodies) vs number theory ; relativity of continuous and discrete. I look at them from the modern perspective of non–commutative geometry, inspired by several deep insights of Alain Connes :

(a) Connes' bold attack on the Riemann Hypothesis ([11], [12]) ;
(b) the discovery (joint with J. Bost, [5]) of a statistical system with the symmetry group $Gal\,(\overline{\mathbb{Q}}/\mathbb{Q})$ and spontaneous symmetry breaking at the pole $s=1$ of the Riemann zeta ;
(c) Connes' idea that approximately finite dimensional simple central C^*–algebras will furnish the missing nontrivial Brauer theory for archimedean arithmetic infinity ([12], p. 38).

I discuss some of these and related themes and add to them the fourth suggestion :

(d) Real Multiplication project, in which elliptic curves with complex multiplication are replaced by two–dimensional quantum tori T_θ whose K_0–group (or rather, its image wrt the normalized trace map $\mathbb{Z}+\mathbb{Z}\theta$) is a subgroup of a real quadratic field. More details will be given in a paper in preparation [32].

2 Real multiplication : an introduction

Let K be a local or global field in the sense of number theory, that is, a field of algebraic numbers, a field of functions on a curve over $\mathbb{F}_q$, or a completion of one of these fields. Denote by K^{ab} its maximal abelian (separable) extension. Class field theory provides a description of the Galois group $Gal\,(K^{ab}/K)$ and some partial information on its action upon K^{ab}. However, specific generators of K^{ab}, together with exact action of $Gal\,(K^{ab}/K)$ upon them, are not generally known. Hilbert's twelfth problem addresses this question. Stark's conjectures, still unproved in general, provide a partial answer : see [59], [60], [65]. Below we will define Stark's numbers directly for $\mathbb{Q}$ (see 2.1) and real quadratic fields (see 4.1).

The most elementary, complete and satisfactory description is furnished by the Kronecker–Weber theorem which we already mentioned.

The Kronecker–Weber machine (KW) can be successfully imitated for at least three more classes of ground fields K : complex quadratic extensions of $\mathbb{Q}$ (referred to as CM, for complex multiplication theory), finite extensions of $\mathbb{Q}_p$ (LT, for Lubin–Tate formal groups), and global fields of finite characteristics (D, for Drinfeld's modules).

The CM case was elaborated first, and all variations on the CM theme, including KW, can be described according to the following scheme. (Only the Lubin–Tate case slightly diverges from the general pattern at this point, see below).

Let K be a global field as above, $O \subset K$ its appropriate subring with quotient field K. We denote by $\mathbb{A}$ the analytic (or formal) additive group defined over K and construct the quotient group $\mathbb{A}/O$ where O is embedded as a subgroup of K–points of $\mathbb{A}$. By construction, multiplication by O induces on $\mathbb{A}/O$ endomorphisms in an appropriate category of K–groups : this is the *geometric* multiplication.

We construct a function t on $\mathbb{A}/O$ taking algebraic (over K) values at the O–torsion points $\xi \in K/O \subset \mathbb{A}/O$. The action of a Galois group upon these torsion points will have an abelian image, if it is shown to commute with the geometric action of O. This presupposes a careful study of the fields of definition of various objects involved.

Class field theory plus special properties of the function t allows us then to establish that the values of t on K/O generate (almost) all of K^{ab}/K. Here are some more details.

2.1 Case (KW).

Here $K := \mathbb{Q}$, $O := \mathbb{Z}$. For $\mathbb{A}$ we will take the additive group considered as a complex–analytic space supplied by an additive coordinate z. The map $a : z \mapsto e^{2\pi iz}$ identifies $\mathbb{A}/O = \mathbb{G}_a^{an}/\mathbb{Z}$ with the (analytic) multiplicative group $\mathbb{G}_m^{an}$. The latter, of course, has a canonical algebraic structure, with a multiplicative coordinate which we will denote $1 + t$, $a^*(1 + t) = e^{2\pi iz}$, and this t will be the function mentioned in the general description. The lattice $\mathbb{Z}$ acts upon $\mathbb{G}_m$ by $1 + t \mapsto (1 + t)^n$: this is the geometric action. The values of $1 + t$ at the torsion points are all roots of unity. The Galois group commutes with geometric semigroup and therefore can be identified with its profinite completion $\widehat{\mathbb{Z}}^*$.

2.2 Case (CM).

Here we denote the ground field (former K) by L and assume that it is an imaginary quadratic extension of $\mathbb{Q}$. For O we take the ring O_L of all integers in L, and for $\mathbb{A}$ the complex analytic additive group as above. The homomorphism of analytic groups $a : \mathbb{G}_a \to X_L$ with the kernel O_L is the universal covering of the elliptic curve X_L. Over $\mathbb{C}$, its field of algebraic functions is generated by the Weierstrass function $\wp(z, O_L)$ and its derivative. The action of O_L upon X_L, in particular, produces rational expressions for $\wp(\alpha z, O_L)$, $\alpha \in O_L$, in terms of $\wp(z, O_L)$, $\wp'(z, O_L)$.

First description of L^{ab}.

An important difference from the (KW)–case is that elliptic curves, unlike $\mathbb{G}_M$, are not rigid but occur in families. In particular, the isomorphism class of X_L over $\mathbb{C}$ is determined by the invariant $j(O_L) \in \mathbb{C}$. It turns out that $L(j(O_L))$ is exactly the maximal unramified abelian extension of L, i.e. its Hilbert class field.

Over $L(j(O_L))$, the maximal abelian extension of L can be generated by the values of the function t such that $a^*(t) := \wp^{-u}(z, O_L)$, where $2u$ is the order of the group O_L^* of units in O_L. Generally, u=1 ; but $u = 2$ (resp. $u = 3$) for $L = \mathbb{Q}(e^{\pi i/2})$ (resp. $L = \mathbb{Q}(e^{\pi i/2})$).

A more geometric description of t is this : X_L/O_L^* is a projective line, and t is a coordinate on it vanishing at the image of the zero point of X_L.

The appearance of X_L/O_L^* rather than X_L itself in this description might seem a minor matter. However, in the conjectural picture of real multiplication it will acquire a great importance, because O_L^* for *real* quadratic fields is always infinite, and the study of the respective quotient will again require tools of noncommutative geometry.

Second description of L^{ab}.

Instead of using only the curve X_L with complex multiplication by the whole O_L, one can consider the family of all elliptic curves Y_L whose endomorphism ring is an order in O_L. It turns out that their absolute invariants $j(Y_L)$ generate a big subfield of L^{ab} : to obtain all of L^{ab}, it remains however to make an additional 2–extension (of infinite degree). The curves Y_L are all isogenous to X_L, and since the points of finite order of X_L "essentially" generate L^{ab}, this gives an intuitive explanation of why this might be so.

However, the two description differ not only by the details of the calculations. What is important in the second picture, is the implicit appearance of the tower or stack of modular curves, parametrizing all elliptic curves, rather than of only one elliptic curve X_L.

This tower consists of the compactified algebraic models of analytic spaces H/Γ, Γ, where H is the upper complex half–plane, and Γ runs over congruence subgroups of $PSL(2,\mathbb{Z})$.

For a readable review of two approaches, see [64] and Ste.

2.3 Case (D).

Here the ground field is the field of functions $k = \mathbb{F}_q(C)$ on a smooth algebraic curve defined over the finite base field with q elements. We choose a point $\infty \in C(\mathbb{F}_q)$ as a part of the structure and denote by O_k the ring of regular functions on the affine curve $C \setminus \{\infty\}$. For $\mathbb{A}$ we take the additive group considered as an analytic group over the quotient field k_∞ of the completion of $\mathcal{O}_{\mathcal{C},\infty}$.

In this situation we have an analogue of exponential function on $\mathbb{A}$:

$$e(z) := z \prod_{\alpha \in O_L \setminus \{0\}} (1 - \frac{z}{\alpha})$$

This function is entire. It is however additive rather than multiplicative : $e(z_1+z_2) = e(z_1)+e(z_2)$. A large abelian extension of k, as in the KW–case, is generated by the values of $e(\lambda z)$ at the "points of finite O_k–order", where λ is an appropriate analogue of $2\pi i$. The action of O_k on $\mathbb{G}_a$ can be alternatively described by the embedding O_k into the (algebraic) crossed product of k_∞ with the semigroup of non–negative powers of the Frobenius endomorphism $x \mapsto x^p$ acting upon k_∞.

This construction can be considerably generalized by replacing λO_k in the construction of the exponential function with any O_k lattice of rank $d \geq 1$. We get in this way the notion of the Drinfeld module of rank d. The case $d = 1$ produces abelian extensions and is similar both to (KW) and (CM) cases. The rank 2 case looks like a version of the theory of elliptic curves, and an appropriate modification of the general d case produces the Langlands type description of the algebraic closure of K.

A thorough historical and mathematical discussion of the theories (KW), (CM) and (D) can be found in the book [66].

2.4 Case (LT).

Here K is a finite extension of $\mathbb{Q}_p$, $O = O_K$ is the ring of integers in K, and the group G_{LT} denoted formerly by $\mathbb{A}/O$ above, is an one–dimensional formal group over O_K equipped with a homomorphism $O_K \to End\, G_{LT}$ sending any $a \in O_K$ to a formal map with the linear term $x \mapsto ax$.

I do not know whether G_{LT} can be interpreted as a quotient $\mathbb{A}/O$ in an appropriate category, although the theory of logarithmic functions developed in the context of p–adic Hodge–Tate theory indicates that such an interpretation might exist. Lubin and Tate construct G_{LT} directly using an ingenious calculation with formal series providing simultaneously a description of the category of such formal groups.

2.5 Real multiplication (RM) and noncommutative geometry.

The simplest class of fields M for which no direct description of M^{ab} is known consists of real quadratic extensions of $\mathbb{Q}$.

Below we sketch a possible approach to this problem via non–commutative geometry and suggest its relation with the Stark conjectures. The idea is straightforward enough. An elliptic curve with complex multiplication by, say, the maximal order O_L in a complex quadratic field L has a complex analytic model $\mathbb{C}/O_L$, where O_L is considered as lattice in $\mathbb{C}$. Similarly, one could try to imagine a space $\mathbb{R}/O_M$ where O_M is the ring of integers of a real quadratic field M. Of course, O_M is not discrete anymore, but the first principle of the non–commutative philosophy says that one should not shy away from such situations : $\mathbb{R}/O_M$ exists as a non–commutative space represented e.g. by a two–dimensional torus T_θ where $O_M = \mathbb{Z} + \mathbb{Z}\,\theta$. Recall that the C^*–algebra of T_θ is the universal algebra generated by two unitaries U, V with the commutation rule $UV = e^{\pi i \theta} VU$. This algebra is the crossed product of $C(\mathbb{R}/\mathbb{Z})$ and the irrational rotation automorphism of $\mathbb{R}/\mathbb{Z}$ induces by the shift $t \mapsto t + \theta$. Crossed products generally serve to represent quotients with respect to "bad" (and occasionally good) equivalence relations : cf, [8], pp. 85–91 for a very livrly account of this principle. (With a hindsight, we noted that in the construction of Drinfeld modules the central role was played by a universal crossed product of an algebra of power series and he semigroup generated by the Frobenius endomorphism.)

As the next step, we want to make sense of the statement that T_θ admits real multiplication by O_M. This is obvious enough for $\mathbb{R}/O_M$ understood set–theoretically : the action of O_M is simply induced by the multiplication $O_M \times \mathbb{R} \to \mathbb{R}$. But crossed product C^*–algebras are not functorial in any naive sense. One way to deal with this difficulty is to replace the usual homomorphisms of associative rings by Morita functors between the categories of their (say, right) modules. Morita functors are given by bimodules, and composition of functors corresponds to the tensor product of bimodules.

Similarly, a coarse moduli space of two–dimensional quantum tori up to Morita equivalence can be seen as a quotient $PGL(2,\mathbb{Z}) \setminus \mathbb{P}^1(\mathbb{R})$; various rigidities lead to its modular

covers. Basics of topology and function theory of these spaces were studied in [33].

Finally, a key issue for arithmetics is the question of fields of definition of various objects : of the "non–commutative elliptic curves with real multiplication T_θ, " of the Galois–Hopf automorphism algebras of such objects (replacing their points of finite order) etc. An implication is that we need algebraic geometric, finitely or countably generated objects, preferably over $\mathbb{Z}$, from which the functional analytic structures like C^*–algebras can be obtained by extension of the base field and an appropriate completion.

In the next section I will address some of these challenges. I will start it with a discussion of several constructions in which algebraic numbers, C^*–algebras and related noncommutative rings appear in combinations that are likely to be susceptible to generalizations.

3 Noncommutative geometry and arithmetic

Oversimplifying, one can say that in commutative geometry algebraic numbers appear as values of algebraic functions, whereas in noncommutative geometry they appear as values of traces of projections, or more generally values of appropriate states on observables. In both cases, a control of the action of the Galois group is gained, if this action commutes with an action of certain "geometric" endomorphisms, or correspondences, whenever the latter are defined over the ground field.

We will deal with three basic situations :

(A) V. Jones theory of indices of subfactors (cf. [24], [25], [19]) and its further developments.

(B) Bost–Connes "spontaneous symmetry breaking", with the symmetry group $Gal\,(\overline{\mathbb{Q}}/\mathbb{Q})$ (cf. [5]).

(C) V. Drinfeld's embedding of $Gal\,(\overline{\mathbb{Q}}/\mathbb{Q})$ into the Grothendieck–Teichmüller group : cf. [17], [16], [28].

3.1 Jones indices.

Consider first the Temperley–Lieb algebra $TL\,(n+1,\tau)$ which arose in statistical physics. Over the central subring $\mathbb{Z}[\tau]$, it is generated by the idempotents $e_1,\dots,e_n$ satisfying the relations $e_ie_j = e_je_i$ for $|i-j| \geq 2$ and $e_ie_{i\pm1}e_i = \tau e_i$. It has a finite dimension, and is semisimple for generic τ. For a critical value of the form $\tau^{-1} = 4\,cos^2\frac{m\pi}{n+1}$, $TL\,(n+1,\tau)$ fails to be semisimple.

It is interesting that Stark's numbers for the ground field $\mathbb{Q}$ ([65], p. 79) are differ from these critical numbers only by a sign change and shift by 4 :

$$exp\,(2\zeta'_{(m,n)}(0)) = 4\,sin^2\frac{m\pi}{n}$$

where

$$\zeta_{(m,n)}(s) := \sum_{k\in m+n\mathbb{Z}} |k|^{-s}.$$

In particular, TL–critical values and Stark numbers generate the maximal real subfield of $\mathbb{Q}^{ab}$ and are stable with respect to the Galois group of $\overline{\mathbb{Q}}$.

One can extend the base ring of TL–algebras to $\mathbb{C}$, define the *–involution on them by $e_i^* = e_i$ (so that e_i become projections), and then pass to the inductive limit with respect to the obvious injections $TL\,(n+1,\tau) \mapsto TL\,(n+2,\tau)$. It turns out that the resulting algebra admits an involutive representation in a complex Hilbert space iff either $\tau^{-1} = 4\,cos^2\frac{\pi}{n}$ for some $n \geq 3$, or $\tau^{-1} \geq 4$.

This statement constitutes an essential part of the famous result due to V. Jones who proved that for any pair of II_1–factors $N \subset M$ the index $[M:N]$ lies in the set $\{4\,cos^2\frac{\pi}{n},\ n \geq 3\} \cup [4,\infty]$, and that for the hyperfinite M this is exactly the set of values of indices. Index itself can be defined as a value of the Hattori–Stallings rank, or as a measure of the growth rate of the minimal number of generators of the N–bimodule $M^{\otimes n} := M \otimes_N \cdots \otimes_N M$ as $n \to \infty$.

Notice that although the discrete part of the values of the index constitutes a part of the TL–spectrum, it is not $Gal\,(\mathbb{Q}^{ab}/\mathbb{Q})$–invariant. The reason is that by passing to the C^*–limit, we have implicitly chosen a non–archimedean valuation of $\mathbb{Q}^{ab}$.

3.2 Bost–Connes spontaneous symmetry breaking.

In the remarkable paper [5] the action of $Gal\,(\mathbb{Q}^{ab}/\mathbb{Q})$ on roots of unity appears in yet another setting. Instead of (the union of) the Temperley–Lieb algebras, consider the Hecke algebra $\mathcal{H}$ with involution over $\mathbb{Q}$ given by the following presentation. The generators are denoted μ_n, $n \in \mathbb{Z}_+$, and $e(\gamma)$, $\gamma \in \mathbb{Q}/\mathbb{Z}$. The relations are

$$\mu_n^*\mu_n = 1, \quad \mu_{mn} = \mu_m\mu_n, \quad \mu_m^*\mu_n = \mu_n\mu_m^* \ for\ (m,n) = 1;$$

$$e(\gamma)^* = e(-\gamma), \quad e(\gamma_1+\gamma_2) = e(\gamma_1)\,e(\gamma_2);$$

$$e(\gamma)\mu_n = \mu_n e(n\gamma), \quad \mu_n e(\gamma)\mu_n^* = \frac{1}{n}\sum_{n\delta=\gamma} e(\delta).$$

The idèle class group $\widehat{\mathbb{Z}}^*$ of $\mathbb{Q}$ acts upon $\mathcal{H}$ in a very explicit and simple way : on $e(\gamma)$'s the action is induced by the multiplication $\widehat{\mathbb{Z}}^* \times \mathbb{Q}/\mathbb{Z} \to \mathbb{Q}/\mathbb{Z}$, whereas on μ_n's it is identical.

The algebra $\mathcal{H}$ admits an involutive representation ρ in $l^2(\mathbb{Z}_+)$: denoting by $\{\epsilon_k\}$ the standard basis of this space, we have

$$\rho(\mu_n)\,\epsilon_k = \epsilon_{nk}, \quad \rho(e(\gamma))\epsilon_k = e^{2\pi i k\gamma}\epsilon_k.$$

From this, one can produce the whole $Gal\,(\mathbb{Q}^{ab}/\mathbb{Q})$–orbit $\{\rho_g\}$ of such representations, applying $g \in Gal\,(\mathbb{Q}^{ab}/\mathbb{Q})$ to all roots of unity occuring at the right hand sides of the expressions for $\rho(e(\gamma))\epsilon_k$. All these representations can be canonically extended to the C^*–algebra completion C of $\mathcal{H}$ constructed from the regular representation of $\mathcal{H}$. Let us denote them by the same symbol ρ_g.

To formulate the main theorem of [5], we need some more explanations. The algebra C admits a canonical action of $\mathbb{R}$, which can be interpreted as *time evolution* represented on

the algebra of observables. This is a general (and deep) fact in the theory of C^*–algebras, but for C the action of $\mathbb{R}$ can be quite explicitly described on the generators. Let us denote by σ_t the action of $t \in \mathbb{R}$. *A KMS_β state at inverse temperature β* on (C, σ_t) is defined as a state φ on C such that for any $x, y \in C$ there exists a bounded holomorphic function $F_{x,y}(z)$ defined in the strip $0 \leq Im\, z \leq \beta$ and continuous on the boundary, satisfying

$$\varphi(x\sigma_t(y)) = F_{x,y}(t),\ \ \varphi(\sigma_t(y)x) = F_{x,y}(t + i\beta).$$

Now denote by H the positive operator on $l^2(\mathbb{Z}_+) : H\epsilon_k = (log\, k)\, \epsilon_k$. Then for any $\beta > 1$, $g \in Gal\,(\mathbb{Q}^{ab}/\mathbb{Q})$ one can define a KMS_β state $\varphi_{\beta,g}$ on (C, σ_t) by the following formula :

$$\varphi_{\beta,g}(x) := \zeta(\beta)^{-1}\, Trace\,(\rho_g(x)\, e^{-\beta H}), \quad x \in C$$

where ζ is the Riemann zeta–function. The map $g \mapsto \varphi_{\beta,g}(x)$ is a homeomorphism of $Gal\,(\mathbb{Q}^{ab}/\mathbb{Q})$ with the space of extreme points of the Choquet simplex of all KMS_β states.

To the contrary, for $\beta < 1$ there is a unique KMS_β state. This is a remarkable "arithmetical symmetry breaking" phenomenon.

3.3 $Gal\,(\overline{\mathbb{Q}}/\mathbb{Q})$ and the Grothendieck–Teichmüller group.

Here we briefly describe the setting studied in [17]. Consider the following tower of fields

$$\mathbb{Q}(t) \subset \overline{\mathbb{Q}}(t) \subset F$$

where F is the maximal algebraic extension of $\overline{\mathbb{Q}}(t)$ ramified only at $t = 0, 1, \infty$. The Galois group $Gal\,(F/\overline{\mathbb{Q}}(t))$ is an extension of $\mathcal{G} := Gal\,(\overline{\mathbb{Q}}/\mathbb{Q})$ by $\Pi := Gal\,(F/\overline{\mathbb{Q}}(t))$. Hence $\mathcal{G}$ acts upon Π by outer automorphisms. Now, Π is a profinite completion of the fundamental group $\pi_1(\mathbb{P}^1(\mathbb{C}) \setminus \{0, 1, \infty\})$. Drinfeld prefers to work with the "formal" monodromy group of the differential equation

$$G'(z) = \frac{1}{2\pi i}\left(\frac{A}{x} + \frac{B}{z-1}\right) G(z)$$

where A, B are non–commuting symbols, and with Knizhnik–Zamolodchikov (KZ) scattering data for such an equation treated in a purely algebraic way. This allows him to connect the Galois action with his study of quasi–Hopf algebras.

3.4 Summary.

The three constructions briefly summarized above at the moment do not form a part of a coherent picture. However, they shed some light upon each other.

The Temperley–Lieb algebras were recently understood as a kind of Galois symmetry objects (quantum groupoids) responsible for the classification of pairs of factors $N \subset M$ with finite width and index : see the review [34] and other references quoted therein. Apparently, the action of the cyclotomic Galois group on them comes from a setting similar to that described in [17], which however has to be made explicit yet. Although in the Grothendieck–Drinfeld picture the stress was usually made on the dessins d'enfant rather than unwieldy non–commutative objects, this might be only one of several possibilities.

Approximately finite dimensional algebras are completed inductive limits of split central semisimple algebras, i.e. direct sums of matrix algebras. It would be interesting to develop the theory of inductive limits of general central semisimple algebras over non–closed, in particular, number fields.

No explicit connection was made between Bost–Connes theory and Jones theory. Are there any ? Since zeta function appears explicitly in [5] as a partition function, cyclotomic Stark numbers might surface naturally in this model.

Finally, all three theories are absolute in the sense that their ground field is $\mathbb{Q}$. It is striking that their relative analogs (over arbitrary number fields K) are not known.

To be more precise, some rather straightforward extensions of [5] were studied in [20], [1], [7], [8] but none of them has the π_0 of the idèle class group of K as a symmetry group when $K \neq \mathbb{Q}$. In fact, the archimedean part of the idèles is not properly accommodated, which returns us to the problem of the right Brauer theory for archimedean primes.

One can expect that a real understanding of the classical abelian class field theory will depend on such a generalization.

4 Real multiplication of quantum tori : geometry

4.1 Lattices and pseudolattices.

The category of elliptic curves over $\mathbb{C}$ is equivalent to the category of period lattices. Formally, a lattice is a discrete and cocompact embedding $j : \Lambda \to V$ where Λ is isomorphic to $\mathbb{Z}^2$ and V is an one-dimensional complex vector space ; morphisms are linear maps $V \to V'$ sending $j(\Lambda)$ into $j'(\Lambda')$. The functor establishing equivalence sends (Λ, V, j) to the complex torus $V/j(\Lambda)$.

Let us denote by Λ_τ the lattice $\mathbb{Z}+\mathbb{Z}\tau$, $\tau \in \mathbf{C} \setminus \mathbb{R}$. Then any lattice is isomorphic to some Λ_τ, and each non–zero morphism $\Lambda_{\tau'} \to \Lambda_\tau$ is represented by a non–degenerate matrix

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2, \mathbb{Z})$$

such that

$$\tau' = \frac{a\tau + b}{c\tau + d},$$

It is an isomorphism, iff $g \in GL(2, \mathbb{Z})$. Thus the coarse moduli space classifying lattices up to an isomorphism is

$$PGL(2, \mathbb{Z}) \setminus (\mathbb{P}^1(\mathbb{C}) \setminus \mathbb{P}^1(\mathbb{R}))$$

which is usually written as $PSL(2, \mathbb{Z}) \setminus H$ where H is the upper half–plane.

The map j can degenerate so that j acquires a kernel of rank one ; the image remains discrete but is not cocompact. The coarse moduli space admitting such degenerations is $PSL(2, \mathbb{Z}) \setminus (H \cup \mathbb{P}^1(\mathbb{Q}))$: a cusp is added. This cusp (or several cusps if a rigidity is imposed) can be seen on algebraic geometric models of the modular curves and canonical

elliptic fibrations over them : they essentially correspond to the degenerations of elliptic curves with (stably) multiplicative connected component of the closed fiber.

Classes of the irrational real points constitute another part of the boundary $PGL(2,\mathbb{Z})\setminus \mathbb{P}^1(\mathbb{R})$. This part classifies degenerations of lattices which I will call *pseudolattices* : a pseudolattice is an embedding $j: L \to V$ where L is isomorphic to $\mathbb{Z}^2$, V is an one–dimensional $\mathbb{C}$–space as above, and the closure of $j(V)$ is a real line. Morphisms are defined exactly as for lattices. Denoting by L_θ the pseudolattice $\mathbb{Z}+\mathbb{Z}\theta$, we see as above that any pseudolattice is isomorphic to an L_θ, and morphisms correspond to the fractional linear transformations of θ's with integral coefficients.

These degenerations are invisible in algebraic geometry. because $V/j(L)$ makes no sense as an algebraic or analytic curve. But the machinery of noncommutative geometry of Connes is designed to deal with such spaces. Choosing L_θ as a representative of the respective isomorphism class, we can naively replace $\mathbb{C}/(\mathbb{Z}+\mathbb{Z}\theta)$ by $\mathbb{C}^*/(e^{2\pi i\theta})$ ("Jacobi uniformization"), and then interpret the last quotient as an "irrational rotation algebra", or two–dimensional quantum torus T_θ. We recall that this torus is (represented by) the universal C^*–algebra A_θ generated by two unitaries U, V with the commutation rule $UV = e^{2\pi i\theta}VU$.

The next task is to define morphisms between these quantum tori, so that we could compare their category with that of pseudolattices. Already isomorphisms present a problem : we want invertible fractional linear transforms to produce isomorphic quantum tori. M. Rieffel's seminal discovery (cf. [37]–[39], [41]) was that to this end we should consider Morita equivalences between appropriate categories of modules as isomorphisms between the tori themselves. Taking this lead, we will formally introduce the general Morita morphisms of associative rings, stressing those traits of the formalism that play a central role in the structure theory of quantum tori (of arbitrary dimension).

For a considerably more sophisticated version of such a theory for von Neumann algebras, see [8], VB, and a report [26]. Physical motivations can be found in [42], [77], [14].

4.2 Morita category and projective modules.

Let A, B be two associative rings. A *Morita morphism* $A \to B$ by definition, is the isomorphism class of a bimodule ${}_AM_B$, which is projective and finitely generated separately as module over A and B.

The composition of morphisms is given by the tensor product ${}_AM_B\otimes_B M'_C$, or ${}_AM\otimes_B M'_C$ for short.

If we associate to ${}_AM_B$ the functor

$$Mod_A \to Mod_B:\ N_A \mapsto N\otimes_A M_B,$$

the composition of functors will be given by the tensor product, and isomorphisms of functors will correspond to the isomorphisms of bimodules.

We imagine an object A of the (opposite) Morita category as a noncommutative space, right A–modules as sheaves on this space, and the tensor multiplication by ${}_AM_B$ as the pull–back functor.

Two bimodules ${}_AM_B$ and ${}_BN_A$ supplied with two bimodule isomorphisms ${}_AM\otimes_B N_A \to {}_AA_A$ and ${}_BN\otimes_A M_B \to {}_BB_B$ define mutually inverse Morita isomorphisms (equivalences) between A and B. The basic example of this kind is furnished by $B = Mat\,(n, A)$, $M = {}_AA^n{}_B$ and $N = {}_BA^n{}_A$.

Projective right A–modules up to isomorphism are exactly ranges of idempotents in various matrix rings $Mat\,(n, A)$ acting from the left upon (column) vector modules A^n.

We prefer to work with all n simultaneously. So we will denote by $\mathcal{MA}$ the ring of infinite matrices $(a_{ij})\, i, j \geq 1$, $a_{ij} \in A$, $a_{ij} = 0$ for $i + j$ big enough (depending on the matrix in question). Similarly, denote by A^∞ the left $\mathcal{MA}$–module of infinite columns (a_i), $i \geq 1$, with coordinates in A, almost all zero.

Denote by Pr_A the category of finitely generated right A–modules. Denote by pr_A the category whose objects are projectors $p \in \mathcal{MA}$, $\sqrt{}^{\in} = \sqrt{}$, whereas morphisms are defined by

$$Hom\,(p, q) := q\,\mathcal{MA}\sqrt{}$$

with the composition induced by the multiplication in $\mathcal{MA}$.

There is a natural functor $pr_A \to Pr_A$ defined on objects by $p \mapsto pA^\infty$. In order to define it on morphisms, we remark that morphisms $pA^\infty \to qA^\infty$ can be naturally described by matrices in the following way. Clearly, pA^∞ contains the columns p_k of p which generate pA^∞ as a right A–module. We can then apply any $\varphi : pA^\infty \to qA^\infty$ to all p_k and arrange the resulting vectors into a matrix $\Phi \in \mathcal{MA}$ with k–th column $\varphi(p_k)$. One checks that $\Phi p = \Phi$, and since also $q\Phi = \Phi$, we have $\Phi \in q\mathcal{MA}\sqrt{}$. Conversely, any such matrix determines a unique morphism $pA^\infty \to qA^\infty$.

Claim

(a) The functor $pr_A \to Pr_A$ described above is an equivalence of categories.
(b) If pA^∞ is isomorphic to qA^∞, then $p - q \in [\mathcal{MA}, \mathcal{MA}]$.

A trace of A is any homomorphism of additive groups $t : A \to G$ vanishing on commutators; by definition, it factors through the universal trace $A \to A/[A, A]$. Combining it with the matrix trace, we get its canonical extension to $\mathcal{MA}$. From the Claim above it follows that $t(p)$ depends only on the isomorphism class of pA^∞. The class $p\,mod\,[A, A]$ is called *the Hattori–Stallins rank* of pA^∞.

We define $K_0(A)$ as the Grothendieck group of Pr_A. If $N_A \in Pr_A$, $[N_A]$ denotes its class in $K_0(A)$. If N_A is the range of an idempotent p and t is a trace, $t(p)$ depends only on $[N_A]$ and is additive on exact triples, hence t becomes a homomorphism of $K_0(A)$ (it is called dimension in the theory of von Neumann algebras).

Assume now that A is endowed with an additive (linear or antilinear in the case of algebras) involution $a \mapsto a^*$, $(ab)^* = b^*a^*$, $a^{**} = a$. It extends to matrix algebras : $(B^*)_{ij} := B^*_{ji}$. similarly, it extends to $A^\infty \to A_\infty$ and $A_\infty \mapsto A^\infty$, compatibly with the module structures.

In such a context, it makes sense to consider only those projective modules that are ranges of *projections*, that is, $*$–invariant idempotents.

4.3 Morita category of two–dimensional quantum tori.

By definition, two–dimensional quantum tori are objects of the category $\mathcal{QT}$ whose morphisms are isomorphism classes of bimodules ${}_AM_B$ corresponding to $*$–invariant projections. The algebra A_θ has a unique trace t_A which is normalized by the condition $t_A(1) = 1$. We can now define a functor K from $\mathcal{QT}$ to the category of pseudolattices $\mathcal{PL}$. On objects, we put :

$$K(T) = (L_A, V_A, j_A, s_A).$$

Here $L_A := K_0(A)$, the K_0–group of the category of right projective A–modules (as above, given by projections in finite matrix algebras over A) ; V_A is the target group of the universal trace on A, that is, the quotient space of A modulo the completed commutator subspace $[A, A]$. Furthermore, $j_A = t_A : K_0(A) \to V_A$ is this universal trace extended to matrix algebras ; its value on the class of a module, as we already explained, is its value at the respective projection. The pseudolattice $K(A_\theta)$ comes in fact with an additional structure which we will call *orientation* : namely, the cone of effective elements in $K_0(A_\theta)$. The information it carries is exactly the choice of an orientation of the real closure of $j_A(L_A)$.

On morphisms, we put

$$K({}_AM_B)([N_A]) := [N \otimes_A M_B].$$

It takes some work to show that this functor is well defined. Clearly, it is compatible with orientations.

Unlike the case of elliptic curves, K is not quite an equivalence of categories. It is essentially surjective on objects and morphisms of oriented pseudolattices conserving orientations. However, it glues together some Morita morphisms. The most clear cut case of this is furnished by Morita equivalences : if $\otimes_A M_B$ and $\otimes_A M'_B$ produce Morita equivalences, and induce the same isomorphism of pseudolattices, these functors differ by an automorphism of the category Mod_A which is induced by an automorphism of the ring A.

Thus, in order to achieve an equivalence of categories, one should count as equivalent many morphisms of noncommutative tori induced by ring homomorphisms, contrary to the intuition educated on affine schemes. In fact, when the trace is unique, all the standard automorphisms of A_θ act trivially on $K_0(A_\theta)$. But in any case, all endomorphisms of pseudolattices keeping orientation are induced by bimoduli, which is a basis of real multiplication to which we finally turn.

4.4 Complex Multiplication and Real Multiplication.

Endomorphisms of a lattice Λ and of the respective elliptic curve form a ring which always contains $\mathbb{Z}$. The situations when it is strictly larger than $\mathbb{Z}$ can be described as follows.

(a) $End\,\Lambda \neq \mathbb{Z}$ iff there exists a complex quadratic subfield K of $\mathbb{C}$ such that Λ is isomorphic to a lattice contained in K.

(b) If this is the case, denote by O_K the ring of integers of K. There exists a unique integer $f \geq 1$ (conductor) such that $End\,\Lambda = \mathbb{Z} + fO_K =: R_f$, and Λ is a projective module of rank 1 over R_f. Every K, f and a projective module over R_f come from a lattice.

(c) If lattices Λ and Λ' have the same K and f, they are isomorphic if and only if their classes in the Picard group $Pic\, R_f$ coincide.

Automorphisms of a lattice generally form a group $\mathbb{Z}_{\not\ni}$ (ψ is multiplication by ± 1.) However, integers of two imaginary quadratic fields obtained by adjoining to $\mathbb{Q}$ a primitive root of unity of degree 4 (resp. 6) furnish examples of lattices with automorphism group of order 4 (resp. 6). Only these two fields produce lattices with such extra symmetries.

Pseudolattices with real multiplication admit a completely similar description.

Endomorphisms of a pseudolattice L (we omit other structures in notation if there is no danger of confusion) form a ring $End\, L$ with composition as multiplication. It contains $\mathbb{Z}$ and comes together with its embedding in $\mathbb{R}$ as $\{a \in \mathbb{R} \mid aj(L) \subset j(L)\}$.

(a′) $End\, L \neq \mathbb{Z}$ iff there exists a real quadratic subfield K of $\mathbb{R}$ such that L is isomorphic to a pseudolattice contained in K.

(b′) If this is the case, we will say that L is an RM pseudolattice. Denote by O_K the ring of integers of K. There exists a unique integer $f \geq 1$ (conductor) such that $End\, L = \mathbb{Z} + fO_K =: R_f$, and L is a projective module of rank 1 over R_f. The module L is endowed with a total ordering.

Every K, f and a ordered projective module over R_f come from a lattice.

(c′) If pseudolattices L and L' have the same K and f, they are isomorphic if and only if their classes in the Picard group $Pic\, R_f$ coincide.

Unlike the case of lattices, the automorphism group of a pseudolattice is always infinite, it is isomorphic to $\mathbb{Z} \times \mathbb{Z}_2$.

4.5 Quantum tori as "limits" of elliptic curves.

Comparison of the relevant geometric categories suggests that two–dimensional quantum tori can be thus considered as limits of elliptic curves. More specifically, take a family of Jacobi parametrized curves $E_\tau = \mathbb{C}/(e^{2\pi i \tau})$ with $Im\, \tau > 0$ and $\tau \to \theta \in \mathbb{R}$. It is then natural to imagine T_θ as a limit of E_τ.

Fixing a Jacobi uniformization of an elliptic curve (or abelian variety of any dimension) as a part of its structure is necessary, for example, in problems connected with mirror symmetry. In such contexts our intuition seemingly provides a sound picture (cf. a similar discussion in [60], pp. 100, 113–114).

However, limitations of this viewpoint become quite apparent if one has no reason to keep a Jacobi uniformization as a part of the structure, and is interested only in the isomorphim classes of elliptic curves, perhaps somewhat rigidified by a choice of a level structure.

In this case one must contemplate the dynamics of the limiting process not on the closed upper half–plane but on a relevant modular curve X. Letting τ tend to θ along a geodesic, we get a parametrized real curve on X which, when θ is irrational, does not tend to any limiting point. This is what can happen.

(a) Let θ be a real quadratic irrationality, θ' its conjugate. Consider the oriented geodesic in H joining θ' to θ. The image of this geodesic on any modular curve X is supported by a closed loop, which we denote $(\theta', \theta)_X$.

(b) Let θ be as above, and let τ tend to θ along an arbitrary geodesic. Then the image of this geodesic on X has $(\theta', \theta)_X$ as a limit cycle (in positive time).

(c) Each closed geodesic on X is the support of a closed loop $(\theta', \theta)_X$. The union of them is dense in X. It is a strange attractor for the geodesic flow in the following sense. Having chosen a sequence of loops $(\theta'_i, \theta_i)_X$, a sequence of integers $n_i \geq 1$, and a sequence of real numbers $\epsilon_i > 0$, $i = 1, 2, \ldots$, one can find an oriented geodesic winding $\geq n_i$ times in the ϵ_i–neighborhood of $(\theta'_i, \theta_i)_X$ for each i, before jumping to the next loop.

Now let us imagine that we have constructed a certain object $R(E_\tau)$ depending on the isomorphism class of E_τ (perhaps, with rigidity). This object can be a number, a function of the lattice, a linear space, a category ... Suppose also that we have constructed a similar object $\mathcal{R}(\mathcal{T}_\theta)$ depending on the isomorphism class of T_θ, and that we want to make sense of the intuitive notion that $\mathcal{R}(\mathcal{T}_\theta)$ is "a limit of $R(E_\tau)$." Since in the most interesting for us case (a) E_τ keeps rotating around the same loop, there are two natural possibilities :

(i) *The object $R(E_\tau)$ actually "does not depend on τ", and $\mathcal{R}(\mathcal{T}_\theta)$ is its constant value. Here independence generally means a canonical identification of different $R(E_\tau)$, e.g. via a version of flat connection defined along the loop.*

(ii) *The object $R(E_\tau)$ does depend on τ, and $\mathcal{R}(\mathcal{T}_\theta)$ is obtained by a kind of integrating or averaging various $R(E_\tau)$ along the loop.*

The second case looks more interesting, however, it is not immediately obvious that such objects occur in nature. Remarkably, they do, and precisely in the context of real multiplication and Stark's conjecture. In fact, this is how we will interpret the beautiful old calculational tricks due to Hecke : see [21], [22], [23], [53]. See also [5] for a similar observation related to what Darmon calls Stark–Heegner points of elliptic curves.

In this section we will only explain the geometric meaning of Hecke's substitution, whereas the (slightly generalized) calculation itself will be treated in the next section.

Let $K \subset \mathbb{R}$ be a real quadratic subfield of $\mathbb{R}$ and $L \subset K$ an RM pseudolattice. From now on, we denote by $l \mapsto l'$ the nontrivial element of the Galois group of $K/\mathbb{Q}$.

For any real t, consider the following subset of $\mathbb{C}$:

$$\Lambda_t = \Lambda_t(L) := \{\lambda_t = \lambda_t(l) := le^{t/2} + il'e^{-t/2} \,|\, l \in L\}$$

Lemma

(a) $\Lambda_t(L)$ is a lattice.

(b) Any isomorphism $a : L_1 \to L$ in the narrow sense induces isomorphisms $\Lambda_t(L_1) \to \Lambda_{t+c}(L)$ where c is a constant depending only on a and t is arbitrary.

(c) The image of the curve $\{\Lambda_t \,|\, t \in \mathbb{R}\}$ on any modular curve is a closed geodesic. The affine coordinate t along this curve is the geodesic length.

Remark

One can try to relate elliptic curves to quantum tori by treating these curves themselves as objects on noncommutative geometry represented by some version of the relevant crossed

product algebra. In the most direct approach, the latter is a completion of the non–unitary toric algebra generated by U, V with $UV = e^{2\pi i\tau}VU$. Representation of such an algebra are in fact closely related to vector bundles on E_τ as was shown in [3] (following [2]). Developing further this approach, one can hope to see better what happens when one passes to the unitary limit $\tau \to \theta$.

5 Stark's numbers for real quadratic fields

In this section we will explain Stark's conjecture for real quadratic fields and slightly generalize Hecke's method of calculation of these numbers. Since it involves integration along the geodesic loops introduced above, we conclude that Stark's numbers have something to do with the respective quantum tori. However, key parts of the picture are still missing.

5.1 Stark's numbers at $s = 0$.

In this section we fix a real quadratic subfield $K \subset \mathbb{R}$. Denote by $l \mapsto l'$ the action of the nontrivial element of the Galois group of K, and by O_K the ring of integers of K, and put $N(l) = ll'$.

Let L be an arbitrary integral ideal of K which, together with its embedding in $\mathbb{R}$ and the induced ordering, will be considered as a pseudolattice.

Choose also an $l_0 \in O_K$ so that the pair (L, l_0) satisfies the following restrictions :
(i) *The ideals* $\mathfrak{b} := (L, l_0)$ *and* $\mathfrak{f} := L\mathfrak{b}^{-1}$ *are coprime.*
(ii) *Let* ε *be a unit of* K *such that* $\varepsilon \equiv 1 \, mod \, \mathfrak{f}$. *Then* $\varepsilon' > 0$.
Put now

$$\zeta(L, l_0, s) := sgn\, l_0' \, N(\mathfrak{b})^s \sum_{l \in L}^{(u)} \frac{sgn\,(l_0 + l)'}{|N(l_0 + l)|^s} \tag{4.1}$$

where (u) at the summation sign means that one should take one representative from each coset $(l_0 + l)\varepsilon$ where ε runs over all units $\equiv 1 \, mod \, \mathfrak{f}$. Notice that $(l_0 + L)\varepsilon = l_0 + L$ precisely for such units.

With this conventions, our $\zeta(L, l_0, s)$ is exactly Stark's function denoted $\zeta(s, \mathfrak{c})$ on the page 65 of [59] : our $\mathfrak{b}, \mathfrak{f}$ have the same meaning in [59], and our l_0 is Stark's γ. The meaning of Stark's $\mathfrak{c}$ is explained below.

The Stark number of (L, l_0) is defined as

$$S_0(L, l_0) := e^{\zeta'(L, l_0, 0)}\,. \tag{4.2}$$

The simplest examples correspond to the cases when $(L, l_0) = (1)$, $\mathfrak{f} = L$, in particular, $l_0 = 1$.

Notice that pseudolattices which are integral ideals have conductor $f = 1$.

5.2 Stark's conjecture for real quadratic fields.

In [59], Stark conjectures that $S_0(L, l_0)$ are algebraic units generating abelian extensions of K. To be more precise, let us first describe an abelian extension M/K associated with (L, l_0) using the classical language of class field theory. (Our M is Stark's K, whereas our K corresponds to Stark's k.)

In 4.1 above we constructed, starting with (L, l_0), the ideals $\mathfrak{f}$ and $\mathfrak{b}$ in O_K. Let $I(\mathfrak{f})$ be the group of fractional ideals of K generated by the prime ideals of K not dividing $\mathfrak{f}$, and $S(\mathfrak{f})$ be its subgroup called the principal ray class modulo $\mathfrak{f}$. Then Artin's reciprocity map identifies $G(\mathfrak{f}) := I(\mathfrak{f})/S(\mathfrak{f})$ with the Galois group of M/K.

Consider all pairs (L, l_0) as above with fixed $\mathfrak{f}$. It is not difficult to establish that on this set, $S_0(L, l_0)$ in fact depends only on the class $\mathfrak{c}$ of $\mathfrak{b}^{-1}$ in $G(\mathfrak{f})$. Denote the respective number $E(\mathfrak{c})$.

Conjecture.

The numbers $E(\mathfrak{c})$ are units belonging to M and generating M over K. If the Artin isomorphism associates with $\mathfrak{c}$ an automorphism σ, we have $E(1)^\sigma = E(\mathfrak{c})$.

(We reproduced here the most optimistic form of the Conjecture 1 on page 65 of [59] involving $m = 1$ and Artin's reciprocity map).

5.3 Hecke's formulas.

In this section we will work out Hecke's approach to the computation of sums of the type (4.1), cf. [22]. It starts with a Mellin transform so that instead of Dirichlet series (4.1) we will be dealing with a version of theta–functions for real quadratic fields. We start with introducing a class of such theta functions more general than strictly needed for dealing with (4.1) (and more general than Hecke's one).

5.4 Theta functions of pseudolattices.

Let $K \subset \mathbb{R}$ be as in 4.1. We choose and fix the following data : a pseudolattice $L \subset K$, two numbers $l_0, m_0 \in K$ and a number $\eta = \eta_0 + i\eta_1 \in \mathbb{C}$. A complex variable v will take values in the upper half plane; $\sqrt{-iv}$ is the branch which is positive on the upper part of the imaginary axis.

Finally, choose an infinite cyclic group U of totally positive units in K such that the following conditions hold :

(a) $u(l_0 + L) = l_0 + L$ for all $u \in U$.

(b) $tr\, ulm_0 \equiv tr\, lm_0 \,mod\, \mathbb{Z}$, $tr\, ul_0m_0 \equiv tr\, l_0m_0 \,mod\, \mathbb{Z}$ for all $l \in L$, $u \in U$, where $tr := tr_{K/\mathbb{Q}}$.

Let $\varepsilon > 1$ be a generator of U.

Put now

$$\Theta^U_{L,\eta}\begin{bmatrix} l_0 \\ m_0 \end{bmatrix}(v) :=$$

$$\sum_{l_0+l\,mod\,U}(\eta_0\,sgn\,(l_0'+l')+\eta_1\,sgn\,(l_0+l))\,e^{2\pi i\,v|(l_0+l)(l_0'+l')|}e^{-2\pi i\,tr\,lm_0}e^{-\pi i\,tr\,l_0m_0}. \quad (4.3)$$

Notation $l_0+l\,mod\,U$ means that we sum over a system of representatives of orbits of U acting upon l_0+L.

Notice that such U always exists, and that if we choose a smaller subgroup $V\subset U$, then

$$\Theta^V_{L,\eta}\left[\begin{array}{c} l_0 \\ m_0 \end{array}\right](v)=[U:V]\,\Theta^U_{L,\eta}\left[\begin{array}{c} l_0 \\ m_0 \end{array}\right](v).$$

In order to relate these thetas to Stark's numbers, consider the function

$$\Theta^U_{L,1}\left[\begin{array}{c} l_0 \\ m_0 \end{array}\right](v)=\sum_{l_0+l\,mod\,U} sgn\,(l_0'+l')\,e^{2\pi i\,v|(l_0+l)(l_0'+l')|}. \quad (4.4)$$

Then we have

$$\sum_{l_0+lmod\,U}\frac{sgn\,(l_0'+l')}{|N(l_0+l)|^s}=\frac{(2\pi)^s}{\Gamma(s)}\int_0^{i\infty}(-iv)^s\Theta^U_{L,1}\left[\begin{array}{c} l_0 \\ m_0 \end{array}\right](v)\,\frac{dv}{v}. \quad (4.5)$$

We will now show that these RM thetas can be obtained by averaging some theta constants (related to the complex lattices) along the closed geodesics described above.

4.5. Theta constants along geodesics. Starting with the same data as in 4.4, we consider first of all the Hecke family of lattices $\Lambda_t=\Lambda_t(L)$ (see 3.5 above). ¿From l_0 which was used to shift L, we will produce a shift of Λ_t :

$$\lambda_{0,t}:=l_0\,e^{t/2}+il_0'\,e^{-t/2}.$$

The number m_0 determines a character of L appearing in (3.3) : $l\mapsto e^{-2\pi i\,tr\,lm_0}$. Similarly, we will produce a character of Λ_t from

$$\mu_{0,t}:=m_0\,e^{t/2}+im_0'\,e^{-t/2}$$

by using the scalar product on $\mathbb{C}$

$$(x\cdot y)=Im\,xy=x_0y_1+x_1y_0 \quad (4.6)$$

where $x=x_0+ix_1$, $y=y_0+iy_1$. Since $l_0,m_0\in L\otimes\mathbb{Q}$, we have similarly $\lambda_{0,t},\mu_{0,t}\in\Lambda_t\otimes\mathbb{Q}$. Omitting t for brevity, we put :

$$\theta_{\Lambda,\eta}\left[\begin{array}{c} \lambda_0 \\ \mu_0 \end{array}\right](v):=\sum_{\lambda\in\Lambda}((\lambda_0+\lambda)\cdot\eta)\,e^{\pi iv|\lambda_0+\lambda|^2}e^{-2\pi i(\lambda\cdot\mu_0)-\pi i(\lambda_0\cdot\mu_0)}. \quad (4.7)$$

The two types of thetas are related by Hecke's averaging formula :

5.5 Proposition

We have

$$\Theta^U_{L,\eta}\begin{bmatrix} l_0 \\ m_0 \end{bmatrix}(v) = \sqrt{-iv}\int_{-\log\varepsilon}^{\log\varepsilon} \theta_{\Lambda_t,\eta}\begin{bmatrix} \lambda_{0,t} \\ \mu_{0,t} \end{bmatrix}(v)\,dt. \tag{4.8}$$

Proof. The following formulas are valid for $Im\, v > 0$:

$$e^{2\pi i\, v|mm'|} = \sqrt{-iv}\,|m'|\int_{-\infty}^{\infty} e^{-t/2}e^{\pi iv(m^2e^t+m'^2e^{-t})}\,dt =$$

$$\sqrt{-iv}\,|m|\int_{-\infty}^{\infty} e^{t/2}e^{\pi iv(m^2e^t+m'^2e^{-t})}\,dt \tag{4.9}$$

(see e. g. [27], pp. 270–271). In the rhs of (4.3), replace the first exponent by the integral expressions (4.9), using the first version at η_0 and the second at η_1. We get :

$$\Theta^U_{L,\eta}\begin{bmatrix} l_0 \\ m_0 \end{bmatrix}(v) =$$

$$\sqrt{-iv}\int_{-\infty}^{\infty}\sum_{l_0+l\bmod U}(\eta_0\,(l'_0+l')\,e^{-t/2}+\eta_1\,(l_0+l)\,e^{t/2})\times$$

$$e^{\pi iv((l_0+l)^2e^t+(l'_0+l')^2e^{-t})}e^{-2\pi itr\,lm_0}e^{-\pi itr\,l_0m_0}\,dt\,. \tag{4.10}$$

In view of (4.6) we have

$$\eta_0\,(l'_0+l')\,e^{-t/2}+\eta_1\,(l_0+l)\,e^{t/2} = ((\lambda_{0,t}+\lambda_t)\cdot\eta),$$

$$(l_0+l)^2e^t+(l'_0+l')^2e^{-t} = |\lambda_{0,t}+\lambda_t|^2,$$

and similarly

$$tr\,lm_0 = (\lambda_t\cdot\mu_{0,t}),\quad tr\,l_0m_0 = (\lambda_{0,t}\cdot\mu_{0,t}).$$

Inserting this into (4.10), we obtain

$$\sqrt{-iv}\int_{-\infty}^{\infty}dt\sum_{l_0+l\bmod U}((\lambda_{0,t}+\lambda_t)\cdot\eta)e^{\pi i\,v|\lambda_{0,t}+\lambda_t|^2}\,e^{-2\pi i\,(\lambda_t\cdot\mu_{0,t})}e^{-\pi i\,(\lambda_{0,t}\cdot\mu_{0,t})}. \tag{4.11}$$

Replacing l_0+l by $\varepsilon(l_0+l)$ is equivalent to replacing t by $t+2\log\varepsilon$. Hence finally the right hand side of (4.11) can be rewritten as

$$\sqrt{-iv}\int_{-\log\varepsilon}^{\log\varepsilon}dt\sum_{\lambda_t\in\Lambda_t}((\lambda_{0,t}+\lambda_t)\cdot\eta)\,e^{\pi i\,v|\lambda_{0,t}+\lambda_t|^2}e^{-2\pi i\,(\lambda_t\cdot\mu_{0,t})-\pi i\,(\lambda_{0,t}\cdot\mu_{0,t})} \tag{4.12}$$

which is the same as (4.8).

We will now apply Poisson formula in order to derive functional equations for Hecke's thetas.

5.6 Poisson formula

Let V be a real vector space, $\widehat{V}$ its dual. We will denote by $(x \cdot y) \in \mathbb{R}$ the scalar product of $x \in V$ and $y \in \widehat{V}$. Choose a lattice (discrete subgroup of finite covolume) $\Lambda \subset V$ and put

$$\Lambda^! := \{\mu \in \widehat{V} \,|\, \forall \lambda \in \Lambda,\, (\lambda \cdot \mu) \in \mathbb{Z}\}. \tag{4.13}$$

Choose also a Haar measure dx on V and define the Fourier transform of a Schwarz function f on V by

$$\widehat{f}(y) := \int_V f(x)\, e^{-2\pi i (x \cdot y)} dx. \tag{4.14}$$

If $f(x)$ in this formula is replaced by $f(x + x_0)\, e^{-2\pi i (x \cdot y_0) - \pi i (x_0 \cdot y_0)}$ for some $x_0 \in V$, $y_0 \in \widehat{V}$, its Fourier transform $\widehat{f}(y)$ gets replaced by $\widehat{f}(y + y_0)\, e^{2\pi i (x_0 \cdot y) + \pi i (x_0 \cdot y_0)}$.

The Poisson formula reads

$$\sum_{\lambda \in \Lambda} f(\lambda) = \frac{1}{\int_{V/\Lambda} dx} \sum_{\mu \in \Lambda^!} \widehat{f}(\mu), \tag{4.15}$$

and for shifted functions as above

$$\sum_{\lambda \in \Lambda} f(\lambda_0 + \lambda)\, e^{-2\pi i (\lambda \cdot \mu_0) - \pi i (\lambda_0 \cdot \mu_0)} = \frac{1}{\int_{V/\Lambda} dx} \sum_{\mu \in \Lambda^!} \widehat{f}(\mu_0 + \mu)\, e^{2\pi i (\lambda_0 \cdot \mu) + \pi i (\lambda_0 \cdot \mu_0)}. \tag{4.16}$$

5.7 Functional equations for θ and Θ

In order to transform (4.12) using the Poisson formula, we put

$$V = \mathbb{C} = \{x_0 + i x_1\},\ \widehat{V} = \mathbb{C} = \{y_0 + i y_1\}, \tag{4.17}$$

and take (4.6) for the scalar product.

Lemma

Let the lattice $\Lambda_t \subset \mathbb{C}$ be the Hecke lattice. Then the dual lattice $\Lambda_t^!$ with respect to the pairing (4.6) has the similar structure

$$\Lambda_t^! = \Lambda_t(M) := \{m e^{t/2} + i m' e^{-t/2} \,|\, m \in M\} \tag{4.18}$$

where we denoted by $M = L^?$ the pseudolattice

$$M := \{m \in K \,|\, \forall l \in L,\, tr_{K/\mathbb{Q}}(l' m) \in \mathbb{Z}.\}.$$

Proof. Denote by Γ the lattice (4.18). For any $\lambda = l e^{t/2} + i l' e^{-t/2} \in \Lambda_t$ and $\mu = m e^{t/2} + i m' e^{-t/2} \in \Gamma$ we have

$$(\lambda \cdot \mu) = Im\, \lambda \mu = l m' + l' m = tr_{K/\mathbb{Q}}(l m'). \tag{4.19}$$

Therefore this scalar product lies in $\mathbb{Z}$ if $m \in M$ so that $\Gamma \subset \Lambda_t^!$. Clearly, then, Γ must be commensurable with $\Lambda_t^!$, so that the right hand side of (4.19) can be used for computing $(\lambda \cdot \mu)$ on the whole $\Lambda_t^!$. This finishes the proof.

For example, $O_K^? = \mathfrak{d}^{-1}$ where $\mathfrak{d}$ is the different. In fact, this is the standard definition of the different.

Now let l_1, l_2 be two generators of the pseudolattice L. Put

$$\Delta(L) := |l_1 l_2' - l_1' l_2|. \tag{4.20}$$

Clearly, this number does not depend on the choice of generators.

Lemma

Let the Haar measure on V be $dx = dx_0\, dx_1$. Choose generators l_1, l_2 of L. Then

$$\int_{V/\Lambda_t} dx = \Delta(L). \tag{4.21}$$

Proof. If Λ_t is generated by ω_1, ω_2, then the volume (4.21) equals

$$|Re\,\omega_1\, Im\,\omega_2 - Re\,\omega_2\, Im\,\omega_1|.$$

Taking

$$\omega_1 = l_1 e^{t/2} + i l_1' e^{-t/2}, \quad \omega_2 = l_2 e^{t/2} + i l_2' e^{-t/2},$$

we get (4.20).

Lemma

The Fourier transform of

$$f_{v,\eta}(x) := (x \cdot \eta)\, e^{\pi i v |x|^2},\ \eta = \eta_0 + i\eta_1 \tag{4.22}$$

equals

$$g_{v,\eta}(y) := \frac{i}{v^2}\,(y \cdot i\bar{\eta})\, e^{-\frac{\pi i}{v}|y|^2} \tag{4.23}$$

Proof. Putting $w = -iv$ we have

$$f_{v,\eta}(x) = (x_0\eta_1 + x_1\eta_0)\, e^{-\pi w(x_0^2 + x_1^2)},$$

so that its Fourier transform by (4.13) and (4.14) is

$$\eta_1 \int_{-\infty}^{\infty} e^{-\pi w x_0^2}\, e^{-2\pi i x_0 y_1}\, x_0\, dx_0 \cdot \int_{-\infty}^{\infty} e^{-\pi w x_1^2}\, e^{-2\pi i x_1 y_0}\, dx_1 +$$

$$\eta_0 \int_{-\infty}^{\infty} e^{-\pi w x_0^2}\, e^{-2\pi i x_0 y_1}\, dx_0 \cdot \int_{-\infty}^{\infty} e^{-\pi w x_1^2}\, e^{-2\pi i x_1 y_0}\, x_1\, dx_1 =$$

$$(\eta_0 y_0 + \eta_1 y_1)\, \frac{1}{iw^2}\, e^{-\pi \frac{y_0^2 + y_1^2}{w}}.$$

This is (4.23).

A functional equation for θ

Let us now write (4.16) for $f = f_{v,\eta}$ and Λ_t :

$$\sum_{\lambda\in\Lambda_t} ((\lambda_{0,t}+\lambda)\cdot\eta)\, e^{\pi i v|\lambda_{0,t}+\lambda|^2} e^{-2\pi i(\lambda\cdot\mu_{0,t})-\pi i(\lambda_{0,t}\cdot\mu_{0,t})} =$$

$$\frac{i}{\Delta(L)\, v^2} \sum_{\mu\in\Lambda_t^!} ((\mu_{0,t}+\mu)\cdot i\bar{\eta})\, e^{-\frac{\pi i}{v}|\mu_0+\mu|^2} e^{2\pi i(\lambda_{0,t}\cdot\mu)+\pi i(\lambda_{0,t}\cdot\mu_{0,t})}.$$

In the notation (4.7) this means :

$$\theta_{\Lambda_t,\eta}\left[\begin{array}{c}\lambda_{0,t}\\ \mu_{0,t}\end{array}\right](v) = \frac{i}{\Delta(L)\, v^2}\, \theta_{\Lambda_t^!, i\bar{\eta}}\left[\begin{array}{c}\mu_{0,t}\\ -\lambda_{0,t}\end{array}\right]\left(-\frac{1}{v}\right). \tag{4.24}$$

We now can establish a functional equation for Θ^U as well :

5.8 Proposition

We have

$$\Theta^U_{L,\eta}\left[\begin{array}{c}l_0\\ m_0\end{array}\right](v) = \frac{1}{\Delta(L)\, v}\, \Theta^U_{L^?, i\bar{\eta}}\left[\begin{array}{c} m_0 - l_0\end{array}\right]\left(-\frac{1}{v}\right). \tag{4.25}$$

Proof. This is a straightforward consequence of (4.8) and (4.24).

Bibliographie

[1] J. Arledge, M. Laca, I. Raeburn. *Crossed products by semigroups of endomorphisms and the Toeplitz algebras of ordered groups.* Documenta Math., 2 (1997), 115–138.

[2] V. Baranovsky, V. Ginzburg. *Conjugacy classes in loop groups and G–bundles on elliptic curves.* Int. Math. Res. Notices, 15 (1996), 733–751.

[3] V. Baranovsky, S. Evens, V. Ginzburg. *Representations of quantum tori and double–affine Hecke algebras.* e–Print math.RT/0005024

[4] F. Boca. *Projections in rotation algebras and theta functions.* Comm. Math. Phys., 202 (1999), 325–357.

[5] J. Bost, A. Connes. *Hecke algebras, type III factors and phase transitions with spontaneous symmetry breaking in number theory.* Selecta Math., 3 (1995), 411–457.

[6] L. Brown, P. Green, M. A. Rieffel. *Stable isomorphism and strong Morita equivalence of C^*–algebras.* Pacific J. Math., 71 (1977), 349–363.

[7] P. Cohen. *A C^*–dynamical system with Dedekind zeta partition function and spontaneous symmetry breaking.* J. de Th. de Nombres de Bordeaux, 11 (1999), 15–30.

[8] P. Cohen. *Quantum statistical mechanics and number theory.* In : Algebraic Geometry : Hirzebruch 70. Contemp. Math., vol 241, AMS, Pridence RA, 1999, 121–128.

[9] A. Connes. *Noncommutative geometry.* Academic Press, 1994.

[10] A. Connes. *C^*–algebras et géométrie différentielle.* C. R. Ac. Sci. Paris, t. 290 (1980), 599–604.

[11] A. Connes. *Trace formula in noncommutative geometry and the zeros of the Riemann zeta function.* Selecta Math., New Ser., 5 (1999), 29–106.

[12] A. Connes. *Noncommutative geometry and the Riemann zeta function.* In : Mathematics : Frontiers and Perspectives, ed. by V. Arnold et al., AMS, 2000, 35–54.

[13] A. Connes. *Noncommutative Geometry Year 2000.* e–Print math.QA/0011193

[14] A. Connes, M. Douglas, A. Schwarz. *Noncommutative geometry and Matrix theory : compactification on tori.* Journ. of High Energy Physics, 2 (1998).

[15] H. Darmon. *Stark–Heegner points over real quadratic fields.* Contemp. Math., 210 (1998), 41–69.

[16] *The Grothendieck Theory of Dessins d'Enfants (Ed. by L. Schneps).* London MS Lect. Note series, 200, Cambridge University Press, 1994.

[17] V. G. Drinfeld. *On quasi–triangular quasi–Hopf algebras and some groups closely associated with* $Gal(\overline{\mathbb{Q}}/\mathbb{Q})$. Algebra and Analysis 2 :4 (1990) ; Leningrad Math. J. 2 :4 (1991), 829–860.

[18] C. F. Gauss . *Zur Kreistheilung.* In : Werke, Band 10, p.4. Georg Olms Verlag, Hildsheim–New York, 1981.

[19] F. Goodman, P. de la Harpe, V. Jones. *Coxeter graphs and towers of algebras.* Springer, 1989.

[20] D. Harari, E. Leichtnam. *Extension du phénomène de brisure spontanée de symétrie de Bost–Connes au cas des corps globaux quelconques.* Selecta Math., 3 (1997), 205–243.

[21] E. Hecke. *Über die Kroneckersche Grenzformel für reelle quadratische Körper und die Klassenzahl relativ–abelscher Körper.* Verhandl. d. Naturforschender Gesell. i. Basel, 28 (1917), 363–373. (Math. Werke, pp. 198–207, Vandenberg & Ruprecht, Göttingen, 1970).

[22] E. Hecke. *Zur Theorie der elliptischen Modulfunktionen.* Math. Annalen, 97 (1926), 210–243. (Math. Werke, pp. 428–460, Vandenberg & Ruprecht, Göttingen, 1970).

[23] G. Herglotz. *Über die Kroneckersche Grenzformel für reelle quadratische Körper I, II.* Berichte über die Verhandl. Sächsischen Akad. der Wiss. zu Leipzig, 75 (1923), 3–14, 31–37.

[24] V. Jones. *Index for subfactors.* Inv. Math., 72 :1 (1983), 1–25.

[25] V. Jones. *Index for subrings of rings.* Contemp. Math. 43, AMS(1985), 181–190.

[26] V. Jones. *Fusion en algèbres de von Neumann et groupes de lacets (d'après A. Wassermann).* Séminaire Bourbaki, no. 800 (Juin 1995), 20 pp.

[27] S. Lang. *Elliptic Functions.* Addison–Wesley, 1973.

[28] P. Lochak, L. Schneps. *A cohomological interpretation of the Grothendieck–Teichmüller group.* Inv. Math., 127 (1997), 571–600.

[29] Yu. Manin. *Quantized theta–functions.* In : Common Trends in Mathematics and Quantum Field Theories (Kyoto, 1990), Progress of Theor. Phys. Supplement, 102 (1990), 219–228.

[30] Yu. Manin. *Mirror symmetry and quantization of abelian varieties.* In : Moduli of Abelian Varieties, ed. by C. Faber et al., Progress in Math., vol. 195, Birkhäuser, 2001, 231–254. e–Print math.AG/0005143

[31] Yu. Manin. *Theta functions, quantum tori and Heisenberg groups.* Lett. in Math. Physics, 56 (2001), 295–320. e–Print math.AG/001119

[32] Yu. Manin. *Real multiplication and noncommutative geometry.* (In preparation).

[33] Yu. Manin, M. Marcolli. *Continued fractions, modular symbols, and non-commutative geometry.* e–Print math.NT/0102006

[34] D. Nikshych, L. Vainerman. *Finite quantum groupoids and their applications.* e–Print math.QA/0006057.

[35] M. A. Rieffel. *Strong Morita equivalence of certain transformation group* C^**–algebras.* Math. Annalen, 222 (1976), 7–23.

[36] M. A. Rieffel. *Von Neumann algebras associated with pairs of lattices in Lie groups.* Math. Ann., 257 (1981), 403–418.

[37] M. A. Rieffel. *C^*–algebras associated with irrational rotations.* Pacific J. Math., 93 (1981), 415–429.

[38] M. A. Rieffel. *The cancellation theorem for projective modules over irrational rotation C^*–algebras.* Proc. Lond. Math. Soc. (3), 47 (1983), 285–303.

[39] M. A. Rieffel. *Projective modules over higher–dimensional non–commutative tori.* Can. J. Math., vol. XL, No. 2 (1988), 257–338.

[40] M. A. Rieffel. *Non–commutative tori — a case study of non–commutative differential manifolds.* In : Cont. Math., 105 (1990), 191–211.

[41] M. A. Rieffel, A. Schwarz. *Morita equivalence of multidimensional non–commutative tori.* Int. J. Math., 10 (1999), 289–299. e–Print math.QA/9803057

[42] A. Schwarz. *Morita equivalence and duality.* Nucl. Phys., B 534 (1998), 720–738.

[43] A. Schwarz. *Theta–functions on non–commutative tori.* e–Print math/0107186

[44] J. P. Serre. *Complex Multiplication.* In : Algebraic Number Fields, ed. by J. Cassels, A. Frölich. Academic Press, NY 1977, 293–296.

[45] Y. Soibelman. *Quantum tori, mirror symmetry and deformation theory.* Lett. in Math. Physics, 56 (2001), 99–125. e–Print math.QA/0011163.

[46] H. M. Stark. *L–functions at $s = 1$. III. Totally real fields and Hilbert's Twelfth Problem.* Adv. Math., 22 (1976), 64–84.

[47] H. M. Stark. *L–functions at $s = 1$. IV. First derivatives at $s = 0$.* Adv. Math., 35 (1980), 197–235.

[48] P. Stevenhagen. *Hilbert's 12th problem, Complex Multiplication and Shimura reciprocity.* In : Class Field Theory – Its Centenary and Prospect. Adv Studies in Pure Math., 30 (2001), 161–176.

[49] J. Tate. *Les conjectures de Stark sur les fonctions L d'Artin en $s = 0$.* Progress in Math., vol. 47, Birkhäuser, 1984.

[50] S. Vladut. *Kronecker's Jugendtraum and modular functions.* Gordon and Breach, 1991.

[51] Y. Watatani. *Index for C^*–subalgebras.* Mem. AMS, vol. 83, Nr. 424, Providence, RA, 1990.

[52] E. Witten. *Overview of K–theory applied to strings.* e–Print hep-th/0007175

[53] D. Zagier. *A Kronecker limit formula for real quadratic field.* Math. Ann., 213 (1975), 153–184.

[54] D. Zagier. *Valeurs des fonctions zeta des corps quadratiques réels aux entiers négatifs.* Astérisque 41–42 (1977), 135–151.

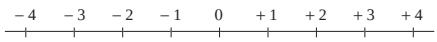
La géométrie des nombres

Christophe Soulé

L'idée première de la géométrie des nombres [1] consiste à considérer les nombres entiers comme contenus dans les nombres réels :

$$\mathbb{Z} \subset \mathbb{R}\,. \tag{1.1}$$

Une telle inclusion peut nous sembler banale, habitués que nous sommes à imaginer les entiers posés sur la droite des réels :



Mais, comme nous le verrons, cela n'est pas toujours allé de soi, d'autant qu'un tel point de vue est très étranger à l'approche algébrique de la théorie des nombres. Or il conduit pourtant à des résultats étonnants, où la géométrie du continu nous fournit des renseignements cruciaux sur des notions arithmétiques discrètes. Nous présenterons ce domaine des mathématiques à travers quelques antécédents historiques et des exemples pris dans les travaux contemporains.

L'approche géométrique de la théorie des nombres est clairement celle que nous propose les Eléments d'*Euclide* [19]. Ce serait un anachronisme que de vouloir chercher dans ce traité l'assertion que les entiers sont contenus dans les réels, puisque n'y apparaissent que les entiers positifs (zéro exclu), et qu'il faudra attendre le dix-neuvième siècle pour que les nombres réels soient définis de façon précise. Mais le livre V des Eléments, celui sur les proportions, joua un rôle décisif dans le genèse des nombres réels chez Dedekind. Et les livres VII à IX exposent les propriétés fondamentales des entiers positifs (nombres premiers, factorisation). Il est intéressant de comparer les premières définitions du livre V et celles du livre VII :

Cinquième livre, Définitions :

1. Une grandeur est partie d'une grandeur, la plus petite de la plus grande, quand la plus petite mesure la plus grande.

2. Une grandeur plus grande est multiple d'une grandeur plus petite, quand la plus grande est mesurée par la plus petite.

3. Une raison est certaine manière d'être de deux grandeurs homogènes entre-elles, suivant la quantité.

Septième livre, Définitions :

1. L'unité est ce selon quoi chacune des choses existantes est dite une.

2. Un nombre est un assemblage formé d'unités.

3. Un nombre est une partie d'un nombre, le plus petit du plus grand, lorsque le plus petit mesure le plus grand.

On notera la similitude entre la définition 1 du livre V et la définition 3 du livre VII. Nous pourrions les traduire en disant qu'une grandeur (resp. un nombre) α "est partie" d'une grandeur (resp. d'un nombre β) si un multiple entier de α est égal à β. Les "raisons" (ou proportions) de la définition 3 du livre V seraient des nombres réels positifs et la situation du livre V est bien décrite par la suite exacte :

$$1 \to R \to G \to U \to 1\,,$$

où R est le groupe multiplicatif des nombres réels positifs (les "raisons"), U est le groupe abélien libre des unités physiques et G est le groupe des "grandeurs". Le monoïde multiplicatif des entiers positifs agit sur G (une grandeur peut en "mesurer" une autre). Mais les entiers ne sont pas vus pour autant comme des raisons (du moins pas à ce stade de la discussion). L'inclusion (1.1), même restreinte au nombres positifs, n'est pas une donnée première de la théorie des proportions d'Euclide .

On notera aussi que les résultats du livre V ne sont pas utilisés dans les preuves des livres VII à IX. Certes, les dessins qui viennent à l'appui des démonstrations dans ces chapitres sont curieusement similaires. Mais l'on n'y trouve pas celui des entiers posés sur la droite des réels, et d'ailleurs beaucoup d'historiens doutent que ces croquis aient figuré dans les versions initiales du traité.

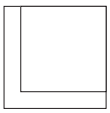
En définitive, c'est surtout par sa conception d'ensemble que le livre des Eléments d'Euclide relie théorie des nombres et géométrie. Ce traité est dans l'ensemble un traité de géométrie et seuls trois livres y parlent des nombres entiers. L'arithmétique occupe, somme toute, une place de second rang.

Après de nombreux travaux à partir de l'œuvre de Diophante, les mathématiciens arabes se sont aussi intéressés à celle d'Euclide. Au dixième siècle, *al-Khujandî* annonce avoir démontré qu'un cube n'est jamais la somme de deux cubes. Il ne nous reste sur cette question que le dessin suivant, dû à Abu Ja'far :

Il s'agit là de la projection sur une de ses faces d'un cube contenant un autre cube. L'auteur déduit de cette figure la formule

$$x^3 - y^3 = x^2(x-y) + (x+y)\,(x-y)\,x$$

d'où il résulte, dit-il, que $x^3 - y^3$ ne peut pas être un cube ! [5] [53].



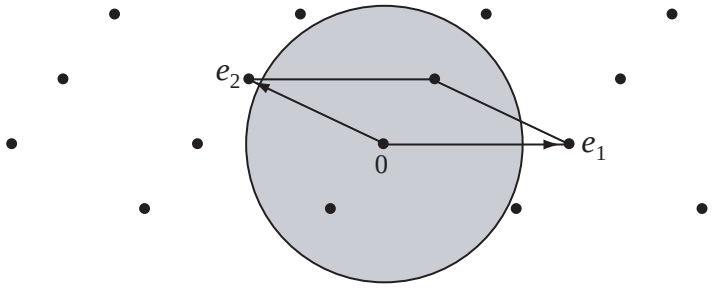
Cet aspect de la préhistoire du théorème de Fermat (on ignore si ce dernier a eu connaissance des travaux ci-dessus) renvoie au thème de notre discussion par le rôle qu'y joue le croquis, bien que celui-ci ne serve qu'à établir une formule algébrique générale. Mais il est aussi intéressant de constater comment les mathématiciens arabes ont su passer des énoncés "bien posés" de Diophante (possédant une solution, en général unique) à un énoncé affirmant l'inexistence d'une solution pour une équation très simple (un énoncé dépourvu d'ailleurs de la moindre conséquence). Notre peur-fascination pour le vide semble être au cœur de notre acharnement à vouloir montrer le théorème de Fermat.

Fermat avait quant à lui une opinion très tranchée (et violente) sur le rapport entre géométrie et arithmétique [5] :

> Il est à peine quelqu'un qui propose des questions purement arithmétiques, il est à peine quelqu'un qui sache les résoudre. Est-ce que l'Arithmétique a plutôt été traitée jusqu'à présent au moyen de la Géométrie que par elle-même ? C'est la tendance qui apparaît dans la plupart des Ouvrages tant anciens que modernes, et dans Diophante lui-même. Car s'il s'est écarté de la Géométrie un peu plus que les autres en astreignant son analyse à ne considérer que les nombres rationnels, il ne s'en est pas dégagé tout à fait, comme le prouvent surabondamment les *Zététiques* de Viète, dans lesquelles la méthode de Diophante est étendue à la quantité continue, et par suite à la Géométrie.
>
> Cependant l'Arithmétique a un domaine qui lui est propre, la théorie des nombres entiers ; cette théorie n'a été que très légèrement ébauchée par Euclide et n'a pas été assez cultivée par ses successeurs (à moins qu'elle n'ait été renfermée dans ces livres de Diophante, dont l'injure du temps nous a privés) ; les arithméticiens ont donc à la développer ou à la renouveler.
>
> Pour éclairer leur marche, je leur propose de démontrer comme théorème ou de résoudre comme problème l'énoncé suivant ; s'ils y parviennent, il reconnaîtront au moins que des questions de ce genre ne le cèdent ni pour la subtilité, ni pour la difficulté, ni pour le mode de démonstration, aux plus célèbres de la Géométrie :
>
> Etant donné un nombre non carré quelconque, il y a une infinité de carrés déterminés tels qu'en ajoutant l'unité au produit de l'un d'eux par le nombre donné, on ait un carré.

Ainsi, s'astreindre à ne considérer que des nombres rationnels ne suffit pas à faire de l'arithmétique, car on risque, ce faisant, de ne trouver que des énoncés également valables pour "la quantité continue" ($\mathbb{Q}$ est dense dans $\mathbb{R}$). L'inclusion (1.1) est dangereuse, elle nous éloigne de l'arithmétique, qui doit s'astreindre à ne traiter que des nombres entiers (positifs). La vraie théorie des nombres n'existe qu'en opposition à la géométrie, et Fermat se propose de faire échapper l'une à l'emprise de l'autre.

L'expression "géométrie des nombres" est due à *Minkowski*; c'est le titre de son ouvrage "Geometrie der Zahlen" (1896). Il y étudie les réseaux euclidiens. On entend par là la donnée d'un sous-groupe discret Λ dans un espace vectoriel réel de dimension finie $\mathbb{R}^n$, tel que le quotient $\mathbb{R}^n/\Lambda$ soit compact.



On a nécessairement $\Lambda \simeq \mathbb{Z}^n$ et l'inclusion $\Lambda \subset \mathbb{R}^n$ est donc une généralisation à n dimensions de l'inclusion (1.1). L'énoncé central de la théorie est le suivant :

Théorème 1. [1] *Il existe une constante $C(n)$ telle que, si le volume de $\mathbb{R}^n/\Lambda$ est inférieur à $C(n)$, le réseau Λ contient un vecteur non nul v de longueur $\|v\|$ inférieure à un.*

Le théorème 1 a un nombre considérable de conséquences arithmétiques. Si F est un corps de nombres, i.e. une extension finie du corps $\mathbb{Q}$ des rationnels, et si $\mathcal{O}_F$ est l'anneau des entiers de F, le Théorème 1 permet de montrer que $\mathcal{O}_F$ n'a qu'un nombre fini de classes d'idéaux et que le groupe $\mathcal{O}_F^*$ de ses éléments inversibles est de type fini. Ce théorème sert aussi à la preuve du théorème d'Hermite selon lequel il n'y a qu'un nombre fini de corps de nombres de degré et de discriminant fixés.

Minkowski donne dans [1] une valeur possible pour la constante $C(n)$. Il remarque aussi que la propriété essentielle de la boule unité qui permet de prouver le Théorème 1 est qu'elle est convexe. Le livre [1] est un des premiers textes sur la géométrie des convexes.

Hilbert tenait en très haute estime cet ouvrage de Minkowski. Il le cite dans trois passages de la présentation de ses fameux problèmes. Il écrit ainsi [23] :

> Mais si je place avant tout la rigueur dans le raisonnement comme condition nécessaire à la solution complète d'un problème, je n'en élèverai pas moins la voix contre cette opinion que ce ne sont que les questions de l'Analyse ou même de l'Arithmétique qui soient seules susceptibles d'un traitement parfaitement rigoureux. Cette opinion émise de temps à autre par des autorités scientifiques, je la regarde comme absolument erronée.
>
> Une notion si étroite de la condition de rigueur conduirait rapidement à ignorer toutes les conceptions tirées de la Géométrie, de la Mécanique et de la Physique.

et

> La coïncidence entre la pensée géométrique et la pensée arithmétique se révèle encore en ceci : dans les recherches arithmétiques, de même que dans les considérations géométriques, nous ne remontons pas à chaque instant la chaîne des déductions jusqu'aux axiomes ; au contraire, lorsque pour la première fois nous attaquons un problème en Arithmétique, exactement comme en Géométrie, nous employons d'abord une combinaison de raisonnements, rapide, inconsciente, non encore définitive, avec une confiance absolue en un certain sentiment arithmétique et en l'efficacité des symboles arithmétiques ; sans cette confiance nous ne pourrions pas plus progresser en Arithmétique que nous ne le pourrions en Géométrie sans la faculté de voir dans l'espace. Comme modèle d'une théorie arithmétique, opérant d'une manière rigoureuse avec les concepts et les symboles de la Géométrie, je citerai l'ouvrage de M. Minkowski : *Geometrie der Zahlen.*

Selon Hilbert, la géométrie peut donc accéder au même niveau de rigueur que l'arithmétique, si elle prend soin d'expliciter ses axiomes. On notera le chemin parcouru depuis Fermat, la théorie des nombres étant maintenant citée en exemple à la géométrie. Plus loin dans [23] (Problème IV, p. 20), Hilbert évoque son article [24] où il cherche à développer axiomatiquement une version non euclidienne de la géométrie de Minkowski.

Au vingtième siècle, la géométrie des nombres a connu des succès considérables. Le Théorème 1 est utilisé dans la plupart des preuves de transcendance et la théorie de l'approximation diophantienne (construction des polynômes auxiliaires). La géométrie des nombres conduit aussi au théorème de *Mordell-Weil* : si A est une variété abélienne sur un corps de nombres F, le groupe de ses point $\Gamma = A(F)$ est de type fini. La preuve de ce résultat (cf. [64]) consiste d'abord à montrer que, pour tout entier $n > 0$, le quotient $\Gamma/n\Gamma$ est fini. On utilise pour cela le théorème d'Hermite cité plus haut. Ensuite, on montre qu'il existe un produit scalaire sur $\Gamma \otimes \mathbb{R}$ ayant de bonnes propriétés (la hauteur de Néron-Tate). Autrement dit, on cherche à montrer que Γ vérifie les propriétés d'un réseau euclidien.

Les années 80 ont été marquées par la preuve due à Faltings, puis Vojta et Bombieri, de la *conjecture de Mordell* : si X est une courbe lisse projective, de genre au moins deux, sur un corps de nombres F, l'ensemble de ses points $X(F)$ est fini. Les preuves de ces trois auteurs utilisent la géométrie des nombres. C'est particulièrement vrai pour celle de Bombieri, qui repose pour l'essentiel sur l'approximation diophantienne et le théorème de Mordell-Weil.

Quelques années plus tôt, *Arakelov* avait proposé une nouvelle approche géométrique des questions diophantiennes [1]. On peut dire – en paraphrasant une formule connue – que, pour Arakelov, "la géométrie des nombres est une géométrie algébrique", et cela en deux sens complémentaires.

D'abord, la géométrie d'Arakelov fournit une synthèse harmonieuse de la théorie des réseaux euclidiens et de celle des schémas de Grothendieck. Si X est un schéma projectif sur $\mathbb{Z}$ et E un fibré algébrique sur X, le groupe des sections globales $\Lambda = \Gamma(X, E)$ de E sur X est un réseau, isomorphe à $\mathbb{Z}^n$ pour un certain entier n. Choisissons une norme sur la fibre E_x de E en chaque point x de la variété complexe $X(\mathbb{C})$. L'espace vectoriel $\Lambda \otimes \mathbb{R}$ est alors muni de la norme sup :

$$\|s\|_{\sup} = \sup_{x \in X(\mathbb{C})} \|s(x)\|$$

si $s \in \Lambda \otimes \mathbb{R}$. La boule unité de $\| \cdot \|_{\sup}$ est convexe et l'on peut appliquer le théorème de Minkowski au couple $(\Lambda, \| \cdot \|_{\sup})$. La géométrie d'Arakelov permet, sous certaines hypothèses, de calculer le covolume de Λ dans $\Lambda \otimes \mathbb{R}$, et donc de produire des sections $s \in \Lambda$ non nulles et de taille bornée [60].

Par ailleurs, on constate que les énoncés de la géométrie des nombres peuvent se formuler dans des termes très voisins de ceux de la géométrie algébrique. En voici un exemple. Si X est une courbe projective et lisse sur un corps k et E un fibré algébrique sur X, le théorème de Riemann-Roch est l'égalité

$$h^0(E) - h^1(E) - \deg(E) = (g-1)\,\mathrm{rg}(E)\,, \tag{1.2}$$

où $h^0(E) = \dim_k \Gamma(X, E)$, $h^1(E) = \dim_k \Gamma(X, E^\vee)$ ($E^\vee$ est le dual de Serre de E), $\deg(E)$ est le degré de E, $\mathrm{rg}(E)$ est son rang, et g désigne le genre de X.

Considérons maintenant un réseau euclidien $\Lambda \subset \mathbb{R}^n$. Notons B la boule unité et

$$\Lambda^\vee = \{v \in \mathbb{R}^n / \forall\, w \in \Lambda\,,\ (v, w) \in \mathbb{Z}\}$$

le dual de Λ. Si $\#\,(\Lambda \cap B)$ est le cardinal de l'ensemble des éléments de Λ contenus dans B, on pose

$$h^0(\Lambda) = \log \#(\Lambda \cap B)$$

et

$$h^1(\Lambda) = \log \#(\Lambda^\vee \cap B)\,.$$

Théorème 2. [10]

$$|h^0(\Lambda) - h^1(\Lambda) + \log(\mathrm{volume}\,(\mathbb{R}^n/\Lambda))| \leq n \log(6) - \log(\mathrm{volume}\,(B))\,. \tag{1.3}$$

Ce théorème ressemble à la formule (1.2) et implique le Théorème 1. En effet, si le volume de $\mathbb{R}^n/\Lambda$ est assez petit, puisque $h^0(\Lambda^\vee) \geq 0$, il résulte de (3) que $h^0(\Lambda)$ est non nul, c'est-à-dire que $\Lambda \cap B$ contient un autre point que l'origine.

Il serait intéressant de reformuler la théorie des réseaux euclidiens dans le cadre de la géométrie non commutative de Connes. On aimerait que les nombres réels $h^0(\Lambda)$ et $h^1(\Lambda)$ (resp. $\log(\mathrm{volume}\,(\mathbb{R}^n/\Lambda))$) soient les dimensions (resp. la dimension virtuelle) de certains modules projectifs sur une algèbre de von Neumann bien choisie.

On notera aussi que (3) est une inégalité alors que (1.2) est une formule exacte. La géométrie des nombres est de nature qualitative. Ainsi le théorème de Mordell-Weil ne dit rien sur le rang de $A(F)$ et l'on ne connaît pas le cardinal de $X(F)$ dans la conjecture de Mordell (bien qu'on sache le majorer). A l'inverse, la théorie algébrique des nombres permet, dans certains cas, de calculer exactement le rang du groupe de Mordell-Weil, le cardinal d'un groupe de classes d'idéaux ou celui d'un groupe de cohomologie.

Comme me l'a fait remarquer L. Clozel, on retrouve cette dichotomie en théorie de Langlands avec, du côté de la géométrie des nombres, les formes automorphes (invariantes par un groupe arithmétique, par exemple le groupe des automorphismes d'un réseau), et, du côté algébrique, la théorie des représentations galoisiennes. Mais peut-on voir l'étude des

formes automorphes comme une sorte de géométrie, au sens (très) élargi que A. Connes donne à ce terme ? C'est un problème ouvert (cf. [16]). Le lecteur trouvera aussi dans [13] une alternative au Théorème 2, qui fait intervenir les séries thêta.

Enfin, on peut rapprocher l'idée d'une distinction entre une arithmétique algébrique exacte et une arithmétique qualitative géométrique des études faites ces dernières années sur le rôle de différentes zones cérébrales dans le traitement élémentaire des nombres entiers. Dans l'article [18] S. Dehaene et ses collaborateurs distinguent une région pariétale et bilatérale où s'effectuerait la comparaison des nombres entiers et leur addition approximative, et une région frontale gauche où se feraient les calculs exacts. La première région est connue pour traiter les images, quand la seconde concerne le langage (ce modèle est affiné dans [15], qui distingue trois régions cérébrales aux lieu de deux). Les méthodes expérimentales utilisées dans ces études sont de trois ordres : l'observation des performances en cas de lésion cérébrale, l'enregistrement de l'activité du cerveau, et la psychologie expérimentale. Par exemple, les auteurs de [18] ont calculé la vitesse nécessaire à l'addition exacte ou approximative de nombres à deux chiffres chez des étudiants russes émigrés aux U.S.A., quand la question était posée dans leur langue maternelle ou en anglais. Ils constatent que le choix de la langue affecte l'addition exacte mais pas l'addition approximative. Ceci conforte l'idée que l'addition exacte fait intervenir les zones du langage (du cerveau frontal gauche) et la "représentation verbale des nombres", mais qu'il n'en est pas de même pour l'addition approximative.

On parle couramment aujourd'hui de "géométrie arithmétique". Mais ce syncrétisme terminologique ne doit pas nous cacher la difficulté, sans cesse renouvelée, du rapprochement des différentes branches des mathématiques. L'arithmétique compte mais elle ne sait voir. La géométrie voit mais elle ne sait parler. L'algèbre parle mais elle ne sait compter.

Bibliographie

[1] M. Minkowski, *Geometrie der Zahlen*, Leipzig, Teubner, 1896.

[2] Euclide, *Les Eléments*, dans *Les œuvres d'Euclide*, traduites par F. Peyrard, A. Blanchard, 1993.

[3] C. Houzel, *Théorème de Fermat, A travers l'histoire de l'analyse diophantienne*, SMF, 1995.

[4] R. Rashed, *Entre arithmétique et algèbre*, Paris, 1984.

[5] P. Fermat, *Second défi aux mathématiciens*, 1657, dans *Œuvres de Pierre Fermat*, I. La théorie des nombres, R. Rashed, C. Houzel, G. Christol 1999, A. Blanchard, pp. 258-259.

[6] D. Hilbert, *Sur les problèmes futurs des mathématiques*, Ed. J. Gabay.

[7] D. Hilbert, *Über die gerade Linie als Kürzeste Verbindung zweier Punkte*, Math. Annalen, **46**, (1895), pp. 91-96.

[8] J.-P. Serre, *Lectures on the Mordell-Weil theorem*, M. Brown Ed., Notes de Michel Waldschmidt. 3ème ed. Aspects of Mathematics. Wiesbaden : Vieweg 1997.

[9] S.Ju. Arakelov, *An intersection theory for divisors on an arithmetic surface*, Izv. Akad. Nauk SSSR, **38**, 1974, pp. 1179-1192.

[10] H. Gillet, C. Soulé, *On the number of lattice points in convex symmetric bodies and their duals*, Israel Journal of Math., **74**, 2/3 1991, pp.347-357.

[11] C. Soulé, D.Abramovich, J.-F.Burnol et J.Kramer : *Lectures on Arakelov geometry*, Cambridge Studies in Advanced Mathematics, **33**, Cambridge University Press 1992.

[12] A. Connes, *Trace formula in noncommutative geometry and the zeroes of the Riemann zeta function*, Selecta Math., **5**, 1999, pp. 29-106.

[13] G. van der Geer, R. Schoof, *Effectivity of Arakelov divisors and the theta divisor of a number field*, Selecta Math., **6**, 2000, pp. 377-398.

[14] S. Dehaene, E. Spelke, P. Pinel, R. Stanescu, S. Tsivkin, *Sources of mathematical thinking : behavioural and brain-imaging evidence*, Science, **284**, 1999, pp. 970-974.

[15] F. Chorbon, L. Cohen, P.F. van de Moortele, S. Dehaene, *Differential contributions of the left and right parietal lobules to number processing*, J. of Cognitive Neuroscience, **11**, 1999, pp. 617-630.

Géométrie complexe et positivité

Pierre Lelong

1 Introduction

Cet exposé décrit la formation d'une géométrie complexe née des problèmes posés par l'étude des fonctions holomorphes, c'est à dire analytiques, de n variables complexes. Le passage de 1 à $n > 1$ pour le nombre des variables complexes a été un révélateur de notions et de problèmes nouveaux, dont l'analyse complexe, en se limitant à une seule variable complexe et à l'image du corps complexe $\mathbb{C}$ donnée sur le plan $\mathbb{R}^2$, ne pouvait aisément donner une idée ni apporter une image sensible.

Les premières recherches s'efforcèrent de caractériser dans les espaces complexes $\mathbb{C}^n$, pour $n > 1$, les domaines d'existence des fonctions holomorphes (domaines d'holomorphie). Ce problème était nouveau. Pour $n = 1$, en effet, l'équivalence topologique de deux domaines entrane dans le plan complexe l'existence d'une application holomorphe (dite conforme) de l'un à l'autre. Cette simplicité disparait pour $n > 1$. On ouvrait donc la voie à des faits insouponnés, voire étranges car le mot analytique associé à une fonction évoquait une régularité. On découvrait dans les espaces complexes une géométrie inconnue où l'analyticité, au contraire, posait des problèmes d'une nature nouvelle auxquels il fallait donner réponse.

On a été conduit alors à définir une géométrie complexe en dimension $n > 1$ qui introduit des objets mathématiques nouveaux. Une remarque importante doit être faite ici : l'introduction de ces objets ne fait pas référence à la structure holomorphe générale. Mais les classes nouvelles introduites devront, par leur définition, être conservées par les applications $\mathbb{C}$-linéaires faites sur les coordonnées complexes et leurs conjuguées dans l'espace complexe $\mathbb{C}^n$.

L'introduction de ces éléments, en développant une géométrie complexe nouvelle, a révélé une propriété particulière de *positivité*. Elle fut définie dans sa généralité en 1957. Elle généralise la positivité des formes " volumes" qui sont homogènes et de bidimension (n, n) dans $\mathbb{C}^n$. Les formes différentielles positives sont devenues un élément important de la géométrie complexe. Utilisées comme " courants " au sens de G. de Rham et opérant sur les formes différentielles de bidegré complémentaire au leur, elles ont défini la classe des opérateurs positifs utilisés aujourd'hui par la géométrie analytique et la géométrie algébrique en y introduisant la notion de courant positif. Son application a permis de définir l'opérateur

d'intégration (W) des formes différentielles sur un ensemble analytique complexe W ayant des singularités et de montrer qu'il est un courant positif et fermé (W) de support W. Ce résultat simple a confirmé une géométrie complexe indépendante de la dimension complexe et a rapproché positivité et holomorphie. La positivité entraine d'autre part que les courants positifs ont la continuité des mesures. Les courants positifs fermés apportent, comme opérateurs linéaires d'intégration sur les formes différentielles, une généralisation à l'intégration sur les ensembles analytiques complexes ;, la multiplicité en un point d'un ensemble analytique complexe étant généralisée par la densité en ce point de la " mesure trace " portée par le support du courant.

Pour $p = 1$, l'étude des courants $(1,1)$ positifs fermés est liée à celle de la classe des fonctions plurisousharmoniques (Psh) introduite dès 1942 indépendamment par K. Oka et l'auteur. Ces fonctions f sont caractérisées par leur semi-continuité supérieure et la positivité de leur courant dérivé $dd^c f$ pour $f \in Psh$.

On ne pouvait donner ici des démonstration toujours complètes, celles-ci étant parfois tributaires de techniques analytiques et de développements plus longs. Le lecteur les trouvera dans la bibliographie donnée à la fin de l'exposé. En particulier signalons l'ouvrage étendu [6] qui développe de nombreuses applications des notions présentées ici, le but du présent exposé étant de montrer comment est née au cours du siècle précédent une géométrie qui a apporté des notions essentielles à la science et s'est développée avec l'analyse complexe lors du passage de celle-ci à plusieurs variables.

Précisons le plan suivi : la section 2 qui suit l'Introduction donne un bref rappel historique sur les propriétés de pseudo-convexité et l'utilisation de la classe Psh des fonctions plurisouharmoniques pour caractériser les domaines d'holomorphie.

Le concept général de positivité est étudié en section 3. La section 4 précise, pour la classe Psh des fonctions plurisousharmoniques une notion de singularité impropre qui joue un rle essentiel pour l'étude de leur prolongement.

Les courants positifs fermés, leur application à l'intégration, font l'objet de la section 5. Leurs propriétés géométriques sont étudiées en section 6 avec le problème fondamental de l'approximation des objets holomorphes.

La section 7 évoque deux applications à des recherches en cours. L'une concerne l'opérateur de Monge-Ampère et la notion de capacité complexe qui en résulte. L'autre évoque la positivité et l'utilisation des courants positifs fermés pour une géométrie complexe en théorie des nombres. Elle associe en géométrie d'Arakelov un " courant de Green " positif fermé à un cycle arithmétique effectif donné dans l'espace projectif complexe.

Le lecteur trouvera à la fin de l'exposé une bibliographie nécessairement restreinte. Elle permet toutefois de se reporter à des ouvrages tels [6] ou, pour des sujets particuliers, à [12] à [10] et à [4], ainsi qu'à des exposés récents : [9] et [14].

2 Débuts et difficultés d'une géométrie complexe

En 1907, Poincaré avait montré l'impossibilité d'obtenir dans l'espace complexe $\mathbb{C}^2$ une bijection holomorphe entre la boule unité et le domaine formé par le produit dans $\mathbb{C}^2$ de deux disques unité de $\mathbb{C}$. Ses résultats témoignent d'une extraordinaire intuition géométrique. Déjà, en 1898, s'inspirant de la mécanique céleste, il avait pressenti le rle de l'opérateur

laplacien et montré que, dans $\mathbb{C}^2$, la fonction $\log|F|$, pour F holomorphe, est la somme d'une fonction harmonique et d'un potentiel newtonien de masse égale à l'aire $F = 0$, (cf. vol 4 de ses " Œuvres " ou [8], p.388 et [6], chap 3, proposition (3-2-15)).
Un problème s'est imposé à l'attention au début du siècle, celui d'obtenir une caractérisation géométrique des domaines d'holomorphie dans $\mathbb{C}^n$ pour $n > 1$. Soit $Hol(G)$ la classe des fonctions holomorphes dans un domaine G de $\mathbb{C}^n$ et soit bG la frontière de G. On dira que G est un domaine d'holomorphie s'il existe $F \in Hol(G)$ telle que tout point de bG ait un voisinage dans lequel F n'est pas prolongeable hors de G, comme fonction holomorphe. E.E. Levi (1910) montre que dans $\mathbb{C}^n$, pour $n > 1$, les domaines d'holomorphie ont une propriété particulière de " pseudo-convexité " qu'on peut définir sur leur frontière si celle-ci est régulière. Mais le problème demeurait de caractériser leur classe (M) dans $\mathbb{C}^n$ par une propriété nécessaire et suffisante. Précisons ici la propriété qui a donné la solution. Une fonction $f : G \to [-\infty, +\infty[$ est dite *exhaustive* dans un domaine G si les ensembles $f(z) < c$ sont relativement compacts dans G. Dès 1906, pour $G \in (M)$, Hartogs avait défini dans G une classe de fonctions d'exhaustion. Mais les fonctions de Hartogs étaient construites à partir de fonctions appartenant à $Hol(G)$. Leur origine excluait d'obtenir par leur classe une propriété géométrique caractéristique des domaines d'holomorphie.

La classe $Psh(G)$ définie à partir des applications linéaires complexes contient la classe des fonctions de Hartogs, elle concide avec elle pour $G \in (M)$, mais elle est strictement plus étendue. La propriété d'exhaustion relativement à Psh, est géométrique et elle est plus précise. Elle a donné la caractérisation cherchée [13]. Pour les difficultés évoquées ici, on pourra se référer à [2]. On notera que le résultat obtenu ne met pas fin à l'étude géométrique de la pseudo-convexité des domaines. Elle intervient en effet dans la solution d'équations fonctionnelles lorsque la solution doit vérifier des conditions à la frontière.

3 La positivité dans les espaces complexes

3.1 Origine et bases du concept

La positivité est présente mais sous une forme particulière dans l'étude de la classe Psh des fonctions plurisousharmoniques. En 1957,dans [11], l'auteur a introduit le concept de positivité complexe sous sa forme générale. Elle s'applique aux formes différentielles homogènes de bidegré (p, p) et aux formes-courants définies auparavant par G. De Rham.

Les opérations du corps complexe $\mathbb{C}$ induisent une orientation sur le plan $\mathbb{R}^2$. L'élément de volume, forme volume, est donné par le produit extérieur anticommutatif et s'écrit : $dV = dx \wedge dy = \frac{i}{2}\, dz \wedge d\overline{z}$.

Dans l'espace $\mathbb{C}^n$, $n > 1$, la forme volume relative à un sous-espace complexe de dimension 1, s'écrit : $dV_1(z) = \frac{i}{2} \sum_k dz_k \wedge d\overline{z_k} = \frac{i}{2}\, \partial\overline{\partial}\, |z|^2 = \tau$.

En dimension complexe $p, (0 < p \leq n)$, dans l'espace $\mathbb{C}^n$, la forme volume s'écrira $dV_p = (p!)^{-1}\, \tau^p$. Les formes volumes sont des formes différentielles homogènes de bidegré (p, p). Par un changement de coordonnées $z'_k = g_k(z_1, \ldots, z_n)$, les g_k étant $\mathbb{C}$-dérivables, on aura, pour $p = n$:

$$dV_n[g(z)] = |det\, \frac{\partial g_k}{\partial z_j}\, |^2\, dV_n(z) = A(g)\tau_n$$

où l'on a $A(g) \geq 0$.

Notons que, si en un point de $\mathbb{C}^n$, on opère un changement de variables exprimant les z_k en fonction des nouvelles variables z'_j, on devra, pour une forme différentielle homogène de type (p,p) faire, non seulement un changement de variables dans les coefficients, mais aussi en faire un sur les différentielles dz_k et leurs conjuguées. On introduit alors une multiplication par des formes linéaires des différentielles en exprimant les anciennes différentielles dz_k en fonction des : dz'_j. On est conduit à la définition suivante :

Définition 3.1 *a) Une forme u homogène de bidegré (n,n) dans $\mathbb{C}^n$ est dite positive si l'on a $u = A\tau^n$ et $A \geq 0$.*

b) Une forme u, homogène de bidegré (p,p) , $1 \leq p < n$, est dite positive si pour toute famille de $q = n - p$ formes linéaires λ_k , le produit $u \wedge i\lambda_1 \wedge \overline{\lambda_1} \wedge \ldots \wedge i\lambda_q \wedge \overline{\lambda_q}$ est une (n,n) forme positive.

La définition est géométrique comme le montre l'énoncé suivant :

Théorème 3.2 *Une forme homogène de bidegré (p,p) est positive si et seulement si sa restriction à tout sous-espace complexe M de dimension complexe p est une forme volume $A(M)\ \tau^p$, où l'on a $A(M) \geq 0$. Dans le cas où on a $A(M) = 0$ pour tout M, on dira que la forme est nulle.*

Exemple. Pour $p = 1$, soit $u = i\sum_{j,k} u_{j,k} dz_j \wedge d\overline{z_k}$. La positivité de u équivaut à la semi-positivité de la forme hermitienne $a \rightarrow \sum u_{j,k}\ a_j \overline{a_k}$.

3.2 Formes décomposables

Les formes positives

$$(1) i\lambda_1 \wedge \overline{\lambda_1} \wedge \ldots \wedge i\lambda_q \wedge \overline{\lambda_q}$$

où les λ_k, $1 \leq k \leq q \leq n$ sont des formes linéaires des différentielles dz_i , $1 \leq j \leq n$, jouent un rôle particulier.

Définition 3.3 *Une forme (p,p) sera dite décomposable si elle est somme finie de formes (1), chacune multipliée éventuellement par un nombre positif.*

Toute forme décomposable est positive. La définition [3.1] de la positivité montre que le produit d'une forme positive non nulle par une forme décomposable est une forme positive, eventuellement nulle. Par contre, en général, le produit d'une forme positive par une forme positive ne sera pas positive sauf si l'une d'entre elles est de type (p,p) pour $p = 0,\ 1,\ n-1,\ n$. On énoncera alors :

Théorème 3.4 *Soient s formes positives. Si toutes ces formes sauf une au plus sont décomposables, leur produit est une forme positive ou nulle.*

3.3 Courants positifs

Les définitions précédentes s'appliquent aux formes différentielles (p,p) représentant un opérateur " courant " au sens de De Rham.

On notera $M'(p,p)$ l'espace des courants repésentés par des formes homogènes de bidegré $(n-p,n-p)$. Dans $\mathbb{C}^n$ ils opèrent sur l'espace $M(p,p)$ des formes homogènes de bidegré (p,p) à coefficients dans l'espace $\mathcal{D}$ des fonctions à support compact et indéfiniment dérivables pour la structure réelle sous-jacente de $\mathbb{C}^n$.

Soient $T \in M'(p,p)$ et $u \in M(p,p)$, où : $T = \sum_{I,J} T_{I,J} dz_I \wedge d\overline{z}_J$ et $u = \sum_{K,S} u_{K,S} dz_K \wedge d\overline{z}_S$. Le calcul de $T(u)$ associe le coefficient $T_{I,J}$ au coefficient $u_{K,S}$, où K est le complémentaire de I et S est celui de J. Le résultat $T(u)$ est ainsi donné par une somme de distributions dont chacune est appliquée à un coefficient de u.

Définition 3.5 *Un courant $T \in M'(p,p)$ est dit positif et non nul si l'on a $T(u) \geq 0$ pour les formes u de $M(p,p)$ qui sont positives et décomposables et si on a $T(u) > 0$ pour l'une d'entr'elles au moins.*

Les propriétés suivantes donnent des définitions équivalentes pour la positivité de $T \in M'(p,p)$.

Théorème 3.6 *La distribution $T \wedge \tau_p$ est une mesure positive et elle le demeure par tout changement des coordonnées locales complexes opérées sur le courant T.*

Ces propriétés utilisent l'invariance par le groupe linéaire complexe de la positivité.

Définition 3.7 *Soit T un courant positif de bidegré $(p;p)$. On appelle trace de T la mesure positive $\sigma_T = T \wedge \tau_{n-p}$.*

Les courants positifs sont réels, on a $T = \overline{T}$. Les distributions associées aux coefficients de T sont des mesures complexes et chacune d'elles, $T_{I,J}$, vérifie une majoration $|T_{I,J}| < c(n,p)\sigma_T$, le coefficient $c(n,p)$ ne dépendant que des dimensions n et p.

4 Fonctions plurisousharmoniques (Psh) et positivité

4.1 Propriétés générales et singularités impropres

La classe Psh introduite en 1942 (P. Lelong, K. Oka) intervient d'une manière essentielle dans l'étude des espaces complexes et des propriétés liées à la positivité, $f \to dd^c f$ donnant une application de la classe Psh dans celle des courants positifs fermés. On a, par définition : $dd^c f = \frac{i}{\pi}\, \partial\, \overline{\partial}\, f$.

Rappelons les définitions suivantes :

Définition 4.1 *On a $f \in Sh(G)$ si, f étant semi-continue supérieurement dans G, on a $-\infty \leq f(z) < +\infty$ et si, pour toute boule $B(z,r) \subset G$, on a $f(z) \leq m_s(f,\ z,\ r)$, m_s étant la moyenne de f sur la sphère $S(z,\ r)$.*

Définition 4.2 *Dans un domaine G de l'espace réel $\mathbb{R}^m$, une fonction $f : G \to \mathbb{R}$ est dite harmonique si elle est dérivable et si elle vérifie*

$\Delta f = \sum_1^m \frac{\partial^2 f}{\partial {x_k}^2} = 0$, *pour tout* $x = (x_k) \in G$.

Définition 4.3 *Soit G un domaine de $\mathbb{C}^n$. Une fonction $G \to [-\infty, +\infty[$ appartient à $Psh(G)$ si elle est semi-continue supérieurement et si elle est sousharmonique sur l'intersection $G \cap L$, pour toute droite complexe L.*

Remarque 4.4 *La définition* **[4.3]** *équivaut, en posant $z = (z_1, \ldots, z_n)$ à*
$f(z) \leq \frac{1}{2\pi} \int_0^{2\pi} f(z + ae^{i\theta})\, d\theta$, pour tout disque $D(z_k, a_k) \subset G$, où
$D(z_k, a_k) = [z'_k = z_k + a_k v\ ,\ v \in \mathbb{C},\ 0 \leq |v| \leq 1\ ,\ 1 \leq k \leq n]$.
La constante $-\infty$ appartient à Sh et à Psh. Une fonction $f \in Psh(G)$, G étant connexe et $f(z) \not\equiv -\infty$, peut être la constante $-\infty$ sur certaines composantes connexes de l'intersection $G \cap L$, où L est un sous-espace.
Les propriétés énoncées ont un caractère local : elles sont vérifiées dans G si et seulement si elles sont vérifiées au voisinage de tout point de G.

Proposition 4.5 *Soit $f \in Sh(G)$. Si le domaine G est connexe, f est la constante $-\infty$ ou, sinon, on a $f \in L^1_{loc}(G)$.*

On dira que f est pluriharmonique si f et $-f$ appartiennent à Psh.

Théorème 4.6 *Soit G un domaine de $\mathbb{C}^n$, et soit $f(z)$ vérifiant*
$-\infty \leq f(z) < +\infty$ pour $z \in G$. On a $f \in Psh(G)$ si les conditions suivantes sont réalisées :
a) On a $f \in L^1_{loc}(G)$.
b) Le courant $dd^c f = \frac{i}{\pi}\, \partial\, \overline{\partial}\, f$ est positif ou nul.
c) Pour tout point $z_0 \in G$ et tout $m > f(z_0)$, il existe un voisinage de (z_0) dans G où l'ensemble $f(z) > m$ est de mesure nulle.

Remarque 4.7 *Dans l'énoncé du théorème* **[4.6]**, *on peut remplacer la condition b) par b') ou b") en précisant que ces conditions seront conservées par tout changement linéaire de coordonnées complexes.*
b') $\Delta f \geq 0$ au sens des distributions
b") On a $f \in Sh(G)$

Proposition 4.8 *Soient $u_1, \ldots, u_p \in Psh(G)$ et $g(t_1, \ldots, t_p)$ une fonction convexe $\mathbb{R}^p \to \mathbb{R}$, croissante des t_k, $1 \leq k \leq p$; alors $g(u_1, \ldots, u_p)$ est plurisousharmonique dans G.*

En particulier on a $\sup_k (u_k) \in Psh(G)$, $\log \sum_1^p \exp u_k \in Psh(G)$ et $\log \sum_1^p |F_k(z)|^{\alpha_k} \in Psh(G)$, pour $F_k(z) \in Hol(G)$ et $\alpha_k \geq 0$ car, pour toute $F \in Hol(G)$, on a $\log |F| \in Psh(G)$.

Proposition 4.9 *1) Pour toute suite décroissante $u_j \in Psh(G)$ et G domaine de $\mathbb{C}^n$, on a :*
$u = \lim u_j \in Psh(G)$.
2) Soit $u \in Psh(G)$, et $u \not\equiv -\infty$, G étant un domaine connexe. Soit ρ_ε une famille de noyaux régularisants, $\mathcal{C}^\infty$ pour $\varepsilon > 0$ et
$u_\varepsilon(a) = \int_{B(a,\ \varepsilon)} u(a + \varepsilon x)\, \rho(x)\, d\tau = u * \rho_\varepsilon(x)$.
Alors $u_\varepsilon(x) \in Psh(G_\varepsilon)$ est $\mathcal{C}^\infty$ et l'on a
$\lim_{\varepsilon \to 0} u_\varepsilon = u(x)$ par valeurs décroissantes pour $\varepsilon \searrow 0$.

Corollaire 4.10 *Soit $u_\alpha \in Psh(G)$ une famille localement majorée et soit $u = \sup_\alpha u_\alpha$. Alors $u^*(z)$, régularisée supérieure de u, définie par : $u^*(z) = \limsup u(z')$ pour $z' \to z$ est plurisousharmonique et est égale à $u(z)$ presque partout dans G.*

La technique analytique associe souvent à une classe $A(G)$ de fonctions définies dans un domaine G, une classe A' d'ensembles de singularités dites impropres. On entend par là qu'une fonction $f \in A(G \setminus M)$ se prolonge nécessairement à G si M est impropre. On établit ainsi une géométrie du prolongement pour la classe A.

Définition 4.11 *Un ensemble M est dit pluripolaire dans un domaine G de $\mathbb{C}^n$ si tout point de G a un voisinage connexe V où $M \cap V$ est contenu dans l'ensemble $f(z) = -\infty$ pour une fonction $f \in Psh(V), f$ n'étant pas la constante $-\infty$ dans V. Tout ensemble pluripolaire est de mesure nulle.*

Théorème 4.12 *Tout ensemble fermé M pluripolaire est une singularité impropre pour les fonctions Psh localement majorées dans son voisinage et leur prolongement à M est unique.*

La propriété de prolongement vaut pour $\log|F|, F$ holomorphe et pour la fonction F elle même. Dans un domaine connexe, le complémentaire d'un ensemble pluripolaire fermé est connexe.

4.2 Pseudo-convexité et domaines d'holomorphie

On a évoqué, (cf section **2**), le problème posé par la caractérisation géométrique des domaines d'holomorphie. La positivité intervient à travers leur pseudo-convexité définie à partir de la classe Psh. On se limitera ici à la classe (H) des domaines G qui sont bornés et univalents dans $\mathbb{C}^n$ en complétant la section **2**.

On dira que G est un domaine pseudo-convexe s'il existe dans G une fonction (dite d'exhaustion), soit $f \in Psh(G)$, telle que les ensembles $f(z) < m$ soient relativement compacts dans G pour tout m. L'existence d'une fonction $F \in Hol(G)$ non prolongeable entrane celle d'une fonction d'exhaustion. Tout domaine d'holomorphie est donc pseudo-convexe.

La démonstration de la réciproque est ramenée au cas des domaines strictement pseudo-convexes où l'exhaustion est réalisée par une fonction $f \in Psh$, telle que $f(z) - c|z|^2$, pour une valeur $c > 0$ soit plurisousharmonique. De tels domaines sont localement d'holomorphie, chaque point de la frontière a un voisinage W tel que $W \cap G$ soit un domaine d'holomorphie. En utilisant une représentation intégrale de $F \in Hol$ on montre que si deux domaines d'holomorphie D et D' coincident dans le domaine $a < x_1 < b$, où $z_1 = x_1 + iy_1$, leur réunion est encore un domaine d'holomorphie. Un tel recollement des domaines a permi, par répétition, de passer du local au global pour la propriété et a conduit au résultat cherché [13].

Une autre méthode résout le problème dans un cadre géométrique différent. Pour montrer l'existence d'une fonction $F \in Hol(G)$, on se place pour F dans un espace fonctionnel défini à partir de $f \in Psh(G)$. Passant à la fonction exponentielle on considérera dans ce but, un espace de Hilbert E dont la norme N est définie pour $f \in Psh(G)$ et $g(z) > 0$ par :
$N(F) = \int_G |F(z)|^2 e^{-f(z)} g(z) d\tau(z) < +\infty$.

L'existence d'une fonction F ayant G comme domaine d'holomorphie résulte alors des propriétés de l'espace E construit à partir de $f \in Psh(G)$. Cette méthode hilbertienne sera utilisée dans la section **6**.

5 Courant positif fermé d'intégration

Soit W un ensemble analytique complexe de dimension p. L'intégration sur W d'une forme différentielle (p,p), du fait que W n'a pas une structure de variété en tout point, nécessitait une recherche pour définir l'opérateur d'intégration $[W]$ sur W en toute dimension complexe : $[W]$ sera obtenu par prolongement du courant positif fermé $|W_r|$ à travers les singularités (cf [11]).

Rappelons qu'un ensemble analytique complexe W sur une variété analytique complexe X est un sous-ensemble fermé de X tel que tout point $z \in W$ ait un voisinage $U \subset X$ tel que $U \cap W$ soit l'ensemble des zéros communs à un nombre fini de fonctions $F_j \in Hol(U)$. On note W_r l'ensemble des points dits " réguliers » au voisinage desquels W est une variété et W_s celui des points singuliers ; W est dit irréductible dans un domaine G s'il n'est pas dans G une réunion d'ensembles analytiques. On supposera dans la suite que G est un domaine connexe de $\mathbb{C}^n$. Si W est irréductible, W_r est dense sur W, W_r est connexe et est alors de la même dimension p, $0 < p < n$ en tous ses points ; p sera dit la dimension complexe de W.

Théorème 5.1 *Pour tout ensemble analytique complexe W dans un domaine $G \subset \mathbb{C}^n$, de dimension p, l'opérateur d'intégration $[W]$ sur W, est un courant positif fermé de bidegré $(n-p,\ n-p)$.*

La démonstration a montré d'abord que le courant d'intégration sur W_r a une masse finie, puis, elle majorait cette masse au-voisinage de W_s. Une étude postérieure d'El Mir a permis de simplifier cette approche en précisant le caractère impropre d'une classe de singularités sur W.

Définition 5.2 *Un ensemble $E \in G$ est dit pluripolaire complet dans G si tout point de G a un voisinage U dans lequel existe une fonction $u \in Psh(U)$ localement sommable telle que $E \cap U$ soit l'ensemble $u(z) = -\infty$ dans U.*

Tout sous-ensemble analytique complexe fermé est pluripolaire complet.

Théorème 5.3 [7] *Soit $E \subset G$ un ensemble fermé pluripolaire complet et soit T un courant positif fermé dans $G \setminus E$. Si la masse de T est finie au voisinage de chaque point de E, l'extension triviale de T (obtenue en prolongeant par des mesures nulles sur E les mesures associées aux coefficients de T) est un courant positif fermé.*

Le théorème **[5.3]** précise donc une classe de singularités impropres. Il abrège la preuve du théorème **[5.1]**, dès qu'on a établi que le courant d'intégration sur l'ensemble W_r des points réguliers a une masse localement finie.

Les conséquences géométriques seront données à la section suivante. Donnons des corollaires du théorème **[5.1]**.

Corollaire 5.4 *Soit $W \subset G$ un ensemble analytique de dimension pure p dans $\mathbb{C}^n$ et W_j ses composantes irréductibles dans G. Alors tout courant fermé T de bidegré (q,q) pour $q = n - p$, de support dans W et ayant la continuité des mesures est de la forme $\sum c_j[W_j]$ où les c_j ont des valeurs complexes. De plus T est positif (il est alors décomposable) si et seulement si les c_j sont tous positifs ou nuls.*

Corollaire 5.5 *Soit T un courant de bidegré (q,q), ayant ainsi que dT la continuité des mesures. Si le support de T est contenu dans un ensemble analytique de dimension complexe inférieure à $n - q$, alors T est nul.*

L'énoncé suivant connu sous le nom d'équation de Lelong-Poincaré (cf la référence donnée [section **2**]) est une conséquence des propriétés précédentes :

Proposition 5.6 *Soit F une fonction méromorphe donnée dans un domaine connexe de $\mathbb{C}^n$ et de diviseur $\sum_j m_j Z_j$. Alors* $\log |F|$ *est localement sommable dans G et vérifie l'équation $dd^c \log |F| = \sum_j m_j [Z_j]$ entre courants positifs fermés de bidegré $(1,1)$.*

6 Géométrie des courants positifs fermés

6.1 Mesures positives et nombres (L)

L'opérateur d'intégration $[W]$ sur un ensemble analytique complexe W de dimension p étant un courant positif fermé, un problème se présente : caractériser la classe $[W]$ dans l'ensemble M' des courants positifs fermés de dimension complexe p. Un nombre $\nu(T,a) \geq 0$ sera défini pour a appartenant au support de T. Il deviendra un objet géométrique, la multiplicité algébrique de W en a si l'on a $T = [W]$. La condition $\nu(T,a) \geq c > 0$ en tout point a entrane que le support de T est un ensemble analytique. Rappelons les notations utilisées, $M'(p,p)$ étant l'emsemble des courants de bidimension (p,p). On pose :

$d^c = (2i\pi)^{-1}(\partial - \overline{\partial})$, $dd^c = i\pi^{-1}\partial\overline{\partial}$. Pour $f \in Psh(G)$, le courant $dd^c f$ est fermé et positif dans G. La mesure masse σ de $T \in M'(p,p)$ est définie par $\sigma = T \wedge \tau_p$ où τ_p est l'élément de volume de $\mathbb{C}^p$.

Définition 6.1 *Une fonction $g \in Psh(G)$ est dite semi-exhaustive si $g(z) < m$ est relativement compact dans G pour $m < R < +\infty$.*

On associera à un courant positif fermé des mesures calculées avec un " poids " $g \in Psh(G)$, semi- exhaustif, exp g étant continu. Pour un courant positif fermé $T \in M'(p,\ p)$, on notera $\nu\ (T,g,m)$ la mesure positive calculée avec le poids g sur le compact : $g(z) \leq m < +\infty$. Utilisant les propriétés de g et $dT = 0$, on a, pour le compact $B(m)$ défini par $g(z) \leq m < +\infty$:

$\nu(T,\ g,\ m) = \int_{B(m)} T \wedge (dd^c g)^p = \int_{B(m)} T \wedge (dd^c g')^p$ où $g' = \max(g,s)$ pour $s \leq m$. La formule de Stokes, T étant fermé et de dimension complexe p, montre que le résultat ne dépend pas de $s < m$. Lorsque $m \to -\infty$, $\nu(T,\ g,\ m)$ tend en décroissant vers une limite notée $\nu(T,g)$.

Le nombre $\nu(T,g)$ défini (Lelong 1957) avec un poids $g(z) = \log |z - a|$, soit $\nu(T,\ \log |z-a|)$, est dit nombre de Lelong " usuel " de T au point a. La notion fut généralisée

(1986) par Kiselman (nombre L directionnel) en utilisant le poids $g(z) = \sup_k l_k^{-1} \log |z_k - a_k|$, $l_k > 0$, $1 \leq k \leq n$. Demailly (1987) a généralisé en introduisant un poids $g(z)$ donné par une fonction Psh semi-exhaustive.

Dans la suite on s'intéressera au cas où l'ensemble $g(z) \leq m$ contient, pour $-\infty \leq m < m_0$, un seul point de l'ensemble $g(z) = -\infty$ ce qui est le cas pour les nombres de Lelong (usuel et directionnel).

Définition 6.2 *Le nombre positif $\nu(T, g) = \lim \nu(T, g, m)$ obtenu pour $m \to -\infty$ sera appelé nombre L au point a, si $g(z) = \log |z - a|$. On le notera nombre $L(g)$ dans le cas d'un poids g différent. Dans le premier cas, on écrira $\nu(T, a, \log r)$ au lieu de $\nu(T, g, m)$, la boule $g(z) < m = \log r$ étant alors la boule $B(a, r)$ de $\mathbb{C}^n$.*

La positivité de T a pour conséquence :

Proposition 6.3 *Soit $h(z) = k\,[g(z)]$ où $k(t)$ est une fonction $\mathbb{R} \to \mathbb{R}$ convexe et croissante. Alors on a la relation $\nu(T, h, r) = k'(r)\, \nu(T, g, r)$, où k' est la dérivée à gauche de la fonction $k(t)$.*

L'application pour $k(t) = e^t$ donne une propriété simple du nombre de Lelong usuel.

Théorème 6.4 *Soit T un courant positif fermé de bidegré $(n-p, n-p)$, de dimension p complexe, $p < n$, dans un domaine G de $\mathbb{C}^n$. Alors pour toute boule $B(a, r)$ compacte dans G, $\sigma(T, a, r)$ étant la mesure trace induite par T sur $B(a, r)$, on a : $\sigma(T, a, r) = \nu(T, a, r)\, V(2p, r)$, où $V(2p, r)$ est le volume de la boule $B(r)$ en dimension complexe p. De plus $\nu(T, a, r)$ est fonction croissante de r et a une majoration finie si $B(a, r)$ appartient à un compact K de G. En conséquence le nombre de Lelong $\nu(T, a)$ d'un courant T positif fermé a une majoration finie pour $a \in K \subset\subset G$.*

Corollaire 6.5 *La masse $\sigma(T, a, r)$ portée par la boule compacte $B(a, r)$ étant une fonction semi continue supérieurement de a, il en est de même de $\nu(T, a)$. L'ensemble défini dans G par $\nu(T, z) \geq c$ est fermé et de mesure localement finie en dimension $2p$ dans l'espace $\mathbb{R}^{2p}$.*

Remarque 6.6 *Pour $p = n - 1$, on a localement $T = dd^c f$ et $f \in Psh$. L'énoncé précédent montre que la mesure laplacien Δf vérifie $\sigma = (2\pi)^{-1} \Delta f$, σ ayant une densité en dimension $2n-2$ égale à $\nu(T, a)$. Le rapport entre la masse de σ dans $B(a, r)$ et le volume de cette boule en dimension complexe $n-1$ est une fonction croissante de r.*

Le théorème suivant (Demailly 1982) permet de changer de poids g.

Théorème 6.7 *Soient g et h deux fonctions de $Psh(G)$ continues et semi-exhaustives sur le support d'un courant T positif et fermé de bidimension (p, p). On suppose qu'on a $\limsup \frac{h(z)}{g(z)} = l < +\infty$ pour $z \in \operatorname{supp} T$ et $g(z) \to -\infty$. Alors on a $\nu(T, h) \leq l\, \nu(T, g)$ avec égalité si l est la limite du quotient pour $g(z) \to -\infty$, et $z \in \operatorname{supp} T$.*

Il en résulte que le nombre de Lelong est indépendant du choix des coordonnées complexes locales ce qui établit un résultat fondamental :

Théorème 6.8 [17] *En un point a d'un ensemble analytique complexe W le nombre $\nu([W], a)$ du courant $[W]$ d'intégration est égal à la multiplicité de a sur W.*

L'ensemble W étant de dimension pure p sur une variété G de dimension complexe n, au voisinage de a pris comme origine des coordonnées locales, il existe une décomposition $z = (z', z")$, soit $C^n = C^p(z') \times C^{n-p}(z")$, telle que le germe de W au voisinage de l'origine soit contenu dans un cne $|z"| \leq c|z'|$ avec $c > 0$. Le degré de la projection de W sur $\mathbb{C}^p$ est alors presque partout constant et égal à la multiplicité algébrique s de W à l'origine. D'autre part la masse du courant d'intégration $[W]$ étant donnée par l'aire de W, le nombre $\nu(T, 0)$ à l'origine est égal à s, ce qui établit l'énoncé.

Le nombre directionnel introduit par Kiselman, et donné par le poids $g(z) = \sup_k l_k^{-1} \log |z - a_k|$, $l_k > 0$, $1 \leq k \leq n$, généralise la notion d'indice. Pour $F \in Hol(G)$ s'annulant en $a \in G \subset \mathbb{C}^n$, $n > 1$, la définition de l'indice $I(F, a, l)$ du zéro de F en a équivaut à $I(F, a, l) = (l_1 \ \ldots \ l_n)\, \nu(T, g)$, où $T = dd^c \log |F|$. L'indice s'étend alors à la classe Psh et aux courants positifs fermés (Lelong et Rashkovskii, 1999). L'indice étant utilisé en théorie des nombres, le nombre directionnel s'applique à des problèmes de géométrie arithmétique complexe.

6.2 Géométrie des nombres L et théorème de Siu

On note $\nu(T, z)$ le nombre L usuel d'un courant positif T en un point z.

Proposition 6.9 *Si T est positif fermé et de dimension p dans G, l'ensemble $E(c)$ des points où l'on a $\nu(T, z) \geq c$, pour $c > 0$ est fermé et de mesure de Hausdorff finie en dimension réelle $2p$ dans G.*

D'autre part l'ensemble $E(c)$ a la propriété suivante (Lelong 1976) :

Proposition 6.10 *Etant donné un courant T positif fermé dans un domaine pseudo-convexe borné G de $\mathbb{C}^n$, il existe dans G un courant T' positif fermé de bidegré $(1, 1)$, de bidimension $(n-1,\ n-1)$, qui en tout point z de G vérifie $\nu(T', z) = \nu(T, z)$.*

Cette proposition a simplifié la démonstration de l'important résultat :

Théorème 6.11 *(*[15] *Si T est un courant positif fermé de bidimension (p, p) sur une variété complexe G, les ensembles $E(c, T)$ définis par $\nu(T, z) \geq c$, $c > 0$ sont dans G des sous-ensembles analytiques complexes de dimension complexe p au plus.*

On est ramené au cas $p = n - 1$. On a alors $T = dd^c f$ localement pour $f \in Psh$. Le problème est d'approcher f par des fonctions du type $\log |F|$, F holomorphe, ce qui conduit à étudier l'espace de Hilbert E des fonctions $F \in Hol(G)$ pour lesquelles on a : $\int_G |F(z)|^2\, e^{-f(z)}\, dV(z) < +\infty$.

Le résultat suivant (Hörmander, Bombieri, Skoda) est fondamental, voir [16] :

Théorème 6.12 *Soit $f \in Psh$ où G est un domaine pseudo-convexe de $\mathbb{C}^n$. Alors pour tout $\epsilon > 0$ et pour chaque point a dans G au voisinage duquel e^{-f} est intégrable, il existe $F \in Hol(G)$ vérifiant pour $\epsilon > 0$:*
$F(a) = 1$ et $\int_G |F(z)|^2\, e^{-f(z)}\, [1 + |z|^2]^{-n-\epsilon}\, dV(z) < +\infty$

Pour $f \in Psh(G)$, l'ensemble A des points de G au voisinage desquels e^{-f} n'est pas sommable est un ensemble analytique. En effet en un point $a \in A$ toutes les fonctions F

de l'espace de Hilbert E doivent s'annuler et d'autre part, le théorème **[6.12]** montre que l'ensemble des points où elles s'annulent toutes est distinct de G.

Reste à préciser le nombre L de $f \in Psh$, soit $\nu(f,a)$ pour $a \in A$. On notera $k(f,a)$ le sup des $c \geq 0$ tels que $f(z)\, e^{-2cf}$ est sommable dans un voisinage du point a et on notera $\lambda(f,a) > 0$ son inverse appelé multiplicité d'Arnold. Dans l'espace $\mathbb{C}^{n+1}$ des variables $(Z,\ z), z \in G$, le nombre k est le sup des $c \geq 0$ tels que dans $\mathbb{C}^{n+1}$ le volume défini par $|Z| > 1$, et $\log |Z| + cf(z) < 0$, soit fini pour z appartenant à un voisinage de a. On a :

Théorème 6.13 *La multiplicité d'Arnold $\lambda\ (f,a)$ et le nombre L usuel vérifient en un point a :*

$$\frac{1}{n}\ \nu(f,a) \ \leq\ \lambda(f,a) \ \leq\ \nu(f,a)$$

Le théorème **[6.11]** en résulte alors.

Remarque 6.14 *Les limites données pour le théorème* **[6.13]** *sont strictes. Exemples : pour $f = \log |z|$, $z \in \mathbb{C}^n$, on a à l'origine $\nu(f,0) = 1$ et $k(f,0) = 1$ tandis que pour $f = \log |z_k|$ on a $\nu(f,0) = 1$ et $\lambda(f,o) = \frac{1}{n}$.*

Le nombre de Lelong ne renseigne donc qu'imparfaitement sur la multiplicité d'Arnold, qu'il s'agisse de $f \in Psh$ ou d'une singularité algébrique.
Une conséquence du théorème **[6.10]** est la formule de décomposition suivante :

Théorème 6.15 *Pour tout courant $T \in M'(p,p)$ dans un domaine G, on a une décomposition unique $T = \sum_{j\geq 1} m_j\ [A_j] + R$, où $m_j > 0$. Les $[A_j]$ sont les courants d'intégration sur des ensembles analytiques irréductibles A_J de dimension p et $R \in M'(p,p)$ ne vérifie $\nu(R,z) \geq c > 0$ que sur des ensembles de dimension complexe inférieure à p.*

6.3 Propriétés de prolongement et d'approximation

Nous terminerons cette section en donnant deux résultats récents.

Théorème 6.16 *(Ohsawa-Takegoshi, 1987).*

Soit G un domaine borné et pseudo-convexe de $\mathbb{C}^n$. Soient $f \in Psh(G)$ et X un sous-espace linéaire de dimension complexe p. Alors toute fonction $F \in Hol\ (G \cap X)$ qui vérifie $\int_{G\cap X} |F|^2\ e^{-f}\ \tau_p < +\infty$ se prolonge en une fonction holomorphe dans G avec la majoration

$$\int_G |F|^2\ e^{-f}\ \tau_n \ \leq\ A(p,n,diam.G) \int_{X\cap G} |F|^2\ e^{-f}\ \tau_p$$

D'autre part, en précisant l'emploi des approximations L^2 dans (G), on a :

Théorème 6.17 *(Voir Demailly 1992 dans* [4]*). Dans un domaine G borné pseudo-convexe de $\mathbb{C}^n$, soit $u \in Psh(G)$. On considère pour $m > 0$ l'espace de Hilbert $H(mu,G)$ des fonctions $F \in Hol(G)$ pour lesquelles on a $\int_G |F|^2\ e^{-2mu}\ \tau_n < +\infty$. Soit $s_{i,m}$ une base hilbertienne de l'espace $H(mu,G)$. On pose pour m fixé :*

$$u_m(z) \ =\ \frac{1}{2m}\ \log \sum\ |s_{i,m}(z)|^2 \ \in\ Psh(G).$$

Il existe alors des constantes positives $c_1, c_2 > 0$ *telles qu'on ait pour* $z \in G$ *et tout* $r < dist(z, bG)$ *:*

$$u(z) - \frac{c_1}{m} \leq u_m(z) \leq \sup_{B(r,z)} [u(z)] + \frac{1}{m} \log \frac{c_2}{r^n}$$

En particulier $u_m \to u$ en tout point $x \in G$ pour la topologie $L^1_{loc}(G)$ quand $m \to +\infty$. On a alors :

$$\nu(u, x) - \frac{n}{m} \leq \nu(u_m, x) \leq \nu(u, x)$$

pour $x \in G$.

La multiplicité d'Arnold vérifie

$$\lambda(u, x) - \frac{1}{n} \leq \lambda(u_m, x) \leq \lambda(u, x).$$

Les fonctions u_m donnent également un contrle des nombres L directionnels. Les deux énoncés donnés se prêtent à l'étude dans un cadre assez large (variétés kählériennes, espaces fibrés) des propriétés géométriques liées à la positivité.

7 Les courants positifs et les géométries complexes

Les géométries complexes, aujourd'hui, s'individualisent en fonction de l'univers qu'elles explorent. La positivité intervient alors à travers des propriétés particulières comme les précédentes. Il s'agit là d'un développement qui, au sein d'une science fondamentale, a besoin de regrouper des techniques et de les saisir à travers une géométrie. Nous donnerons deux exemples. L'un concerne le produit de courants positifs. Il définit l'opérateur de Monge-Ampère complexe qui permet l'étude géométrique d'une notion de capacité complexe. L'autre exemple évoque le courant d'intégration en géométrie d'Arakelov, élément d'une géométrie arithmétique. On se limitera aux propriétés de base sans développer les applications à la géométrie des espaces fibrés complexes.

7.1 Produits de courants positifs fermés

Les courants positifs fermés $dd^c\ u$, de bidegré $(1,1)$, pour $u \in Psh$ jouent le rle d'éléments minimaux parmi les courants positifs fermés. D'où l'intérêt du produit uT qui associe $u(z) \in Psh$ à un courant T positif fermé. Si u est dérivable, on a $dd^c\ u \wedge T = dd^c\ (uT)$. Mais on ne peut en général multiplier deux courants ayant la continuité des mesures. Cependant si $|u|$ est borné, on peut définir le produit uT. On définira alors $dd^c u \wedge T$ par la valeur $dd^c\ (uT)$. Cette méthode, généralisée par l'emploi d'une majoration due à Chern, Levine et Nirenberg, a montré l'existence et la commutativité d'un produit de q courants $dd^c\ u_k$, lui-même multiplié par un courant T positif fermé, les $|u_k|$ étant majorés localement. Dans cette méthode, on approche chaque u_k par une suite décroissante de fonctions $u_{k,j} \in Psh$, dérivables, et on passe à la limite du produit au sens de la convergence faible des courants.

Ce résultat, très remarqué (cf Bedford et Taylor [3]), a été suivi d'autres. Ils ont confirmé l'importance des singularités impropres dans l'étude de la positivité complexe. Pour $u \in Psh(G)$,

on notera $L(u)$ le " lieu de u ", ensemble des $z \in G$ tels que la fonction u n'est bornée sur aucun voisinage de z dans G. On obtient :

Théorème 7.1 *(Demailly 1993, Fornaess et Sibony 1995) Soit, dans un domaine G de $\mathbb{C}^n$ un courant T positif fermé de bidimension (p,p) et des fonctions $u_k \in Psh(G)$. Soit L_k le lieu de u_k, $1 \leq k \leq q \leq p$. On note $H(s)$ la mesure de Hausdorff en dimension réelle s. Alors pour $m \leq q$, si l'on a : $H(2p-2m+2)(S) = 0$ pour toute intersection S de $m+1$ ensembles, constituée en prenant le support de T et m des L_k, on a la " propriété du produit " suivante : le produit de T par les q courants $dd^c u_k$ est bien défini comme courant positif fermé. Il a une masse localement finie, et il s'obtient par la convergence faible des courants à partir des suites décroissantes de fonctions Psh régulières qui convergent vers les u_k.*

Le produit de courants du type $dd^c u_k, u_k \in Psh$, par un courant T positif fermé opéré dans les conditions du théorème **7.1** donne donc un courant positif fermé indépendant de l'ordre des facteurs.

7.2 Opérateur de Monge-Ampère et géométrie pluricomplexe

Dans un domaine G de $\mathbb{C}^n$, on appellera opérateur de Monge-Ampère l'opérateur $u \to M(u) = (dd^c u)^n$ qui associe à $u \in Psh(G)$ une mesure positive. Pour $n = 1, M(u)$ définit la distribution $(2\pi)^{-1}\Delta u$, où Δ est l'opérateur laplacien. En dimension complexe $n > 1$ on donnera des représentations " pluripotentielles " de u à partir de $M(u)$. On généralise ainsi le théorème de Dirichlet et les représentations potentielles de la géométrie réelle. Mais dans $\mathbb{C}^n$, pour $n > 1$, l'opérateur de Monge-Ampère et la recherche d'une fonction de Green pour une représentation potentielle de $f \in Psh(G)$ supposent des conditions sur G et sa frontière bG.

Un domaine G de $\mathbb{C}^n$ sera dit hyperconvexe s' il est connexe, et s'il existe $u \in Psh(G)$, continue, vérifiant $u(z) < 0$, les ensembles $u(z) \leq c < 0$ étant compacts dans G. Ce sera le cas si G est pseudo-convexe et bG lipschitzien.

Théorème 7.2 *Si G est hyperconvexe alors, pour $a \in G$, il existe une unique fonction $u(z,a) \in Psh(G)$, continue, nulle sur bG, pour laquelle on a $(dd^c u)^n = (2\pi)^n \delta(a)$, où $\delta(a)$, est la mesure unité en a et où on a : $u(z,a) \sim \log|z-a|$ pour $z \to a$. La fonction $u(z,a)$ est alors l'enveloppe supérieure des $v \in Psh(G)$ qui vérifient $v(z) < 0$ dans G, et $v(z) - \log|z-a|$ borné. La fonction $u(z,a)$ sera appelée la fonction de Green pluricomplexe du domaine G.*

Définition 7.3 *Pour $u(z) \in Psh(G)$, avec $-\infty \leq u(z) < 0$, et $u(z)$ continue et exhaustive, on appellera masse de Monge-Ampère de u, soit $M(G,u)$, la masse de la mesure positive $(dd^c u)^n$ dans le domaine G.*

Théorème 7.4 *Soit $v \in Psh(G)$, continue, exhaustive, de masse totale finie $M(G,v)$ dans G et vérifiant : $-\infty \leq v < 0$. Alors quand r tend vers zéro, la mesure $m(v,r) = (dd^c v)^{n-1} \wedge d^c v$ sur la pseudo-sphère $v(z) = r$, $r < 0$, converge pour $r \to 0$, vers une " mesure au bord " positive m' portée par bG et de masse totale égale à $M(G,v)$. Associée à la fonction de Green, elle donne la représentation cherchée pour toute fonction $f \in Psh(G)$ lorsque f eést continue.*

On renvoie à [5] et à [10], cités dans la bibliographie, pour les calculs et l'utilisation de la formule dite de Lelong-Jensen dans des situations plus générales qui permettent encore de construire un noyau reproduisant les fonctions pluriharmoniques.

7.3 Une notion de capacité pour la géométrie complexe

Soit G un domaine connexe de $\mathbb{C}^n$, soit E un ensemble relativement compact dans G. Soit (E') la famille des fonctions $u \in Psh(G)$ vérifiant $u(z) < 0$ pour $z \in G$ et $u(z) \leq -1$ pour $z \in E$. Soit $w(E,G)(z) = \sup u(z)$, $u \in (E')$ et $v(E,G)(z) = \limsup w(E,G)(z')$ pour $z' \to z$. La fonction v sera dite la régularisée de w dans G.

Théorème 7.5 *On a $v(E,G) \equiv 0$ dans G si et seulement si E est un ensemble G-pluripolaire c'est à dire est contenu dans l'ensemble $f(z) = -\infty$ pour une fonction $f \in Psh(G)$.*

On a $v(E,G)(z) \in Psh(G)$ et on obtient une propriété géométrique du couple (E,G) et de la masse de Monge-Ampère $(dd^c v)^n$ de v.

Théorème 7.6 *Pour un compact $K \subset G$ et G hyperconvexe, on a $[dd^c v(K,G)]^n = 0$ dans l'ouvert $G \backslash K$.*

Définition 7.7 *Soit K un compact dans un domaine G hyperconvexe. On appellera capacité complexe de K dans G la masse de Monge-Ampère $(dd^c v)^n$ de $v = v(K,G)$ portée par K dans G. Elle sera notée $C(K,G)$.*

Définition 7.8 *Un ensemble $A \subset G$ est dit négligeable dans G si l'on a $A = [z \in G : v(z) < v'(z)]$, où v est le* sup *d'une famille de fonctions de $Psh(G)$ négatives et v' sa régularisée supérieure définie par $v'(a) = \limsup v(z)$ pour $z \to a$.*

Théorème 7.9 *Tout ensemble négigeable est pluripolaire.*

Théorème 7.10 *Tout ensemble pluripolaire dans un domaine de $\mathbb{C}^n$ est contenu dans les $-\infty$ d'une fonction $f \in Sh(\mathbb{C}^n)$.*

La notion de capacité complexe appliquée aux compacts et aux ouverts conduit à l'étude des notions de capacité intérieure et de capacité extérieure des ensembles. Les ensembles pluripolaires sont de capacité extérieure nulle. Il en résulte que $f \in Psh(G)$ est bornée dans $G \backslash A$ pour des ensembles A vérifiant $cap\ (A) < \epsilon$, quel que soit $\epsilon > 0$.

7.4 Une généralisation arithmétique

Je terminerai en montrant comment la positivité et la notion de courant positif fermé interviennent dans un domaine différent et contribuent à la formation d'une géométrie arithmétique. Celle-ci s'élabore dans le contexte d'Arakelov (1974), des résultats de Faltings (1991) et de Bost, Gillet et Soulé [1], le point de départ étant donné par une théorie de l'intersection des variétés arithmétiques. Pour une telle variété X dans un espace projectif, l'étude arithmétique définit une hauteur $h(X)$, valeur positive qui, exceptionnellement, peut être nulle lorsque X est un sous-espace de l'espace projectif.

Sur une variété analytique complexe, homogène et compacte X, on associera à un cycle effectif Z un courant positif, " forme de Green ", soit g. Il est dérivable $\mathbb{C}^\infty$ sur $X - [Z]$. Il est localement L^1 sur X et, sur X, il définit un courant dit de Green pour Z, la somme $dd^c g + [Z]$ étant $\mathbb{C}^\infty$ sur X et $[Z]$ étant le courant d'intégration sur le cycle Z.

Des conditions sont toutefois nécessaires sur X si X n'est pas l'espace projectif, ce qui présage des développements ultérieurs.

Dans un espace projectif l'étude donne en particulier un théorème de Bézout. Il majore non plus le degré mais la hauteur $h(X \cdot Y)$ d'une intersection $X \cdot \ Y$ de deux variétés projectives X et Y à partir de leurs degrés et de leurs hauteurs arithmétiques, les degrés étant donnés dans l'espace projectif par celui des fibrés génériques. Sous des conditions assez générales [1] pour le cycle intersection $X \cdot Y$ on obtient la majoration :

$$h(X \cdot Y) \ \leq \ h(X) \mathit{degré}(Y) \ + \ \mathit{degré}(X) h(Y) \ + \ c \ \mathit{degré}(X) \ \mathit{degré}(Y),$$

où $c \ \geq 0$ ne dépend que des dimensions de (X, Y) et est nul si X ou Y est un sous-espace linéaire.

Bibliographie

[1] J.B. Bost, Gillet H., and C. Soulé. Heights of projective varieties and positive Green forms. *J. Amer. Math. Soc.*, 7 :903–1027, 1994.

[2] H. Bremermann. On the conjecture of the equivalence of the plurisubharmonic functions and the Hartogs functions. *Math. Ann.*, 131 :76–86, 1956.

[3] E. Bedford and B.A. Taylor. A new capacity for plurisubharmonic functions. *Acta Math.*, 149 :1–40, 1982.

[4] P. Dolbeault and Al., editors. *International Conference*, volume 188 of *Progress in Math.* Birkh user Verlag, Basel, Boston, Berlin, 2000.

[5] J.P. Demailly. Mesure de Monge-Ampère et mesures pluriharmoniques. *Math. Zeit.*, 194 :519–564, 1987.

[6] J.P. Demailly. Analytic geometry. available at : http ://www-fourier.ujf-grenoble.fr, 1995.

[7] H. El Mir. Sur le prolongement des courants positifs fermés. *Acta Math.*, 153 :1–45, 1984.

[8] P. Griffiths and G. Harris. *Principles of algebraic geometry.* John Wiley, New York, 1978.

[9] C.O. Kiselman. Plurisubharmonic functions and potential theory in several complex variables. Development of Mathematics 1950-2000, 2001.

[10] M Klimek. Pluripotential theory. *Oxford Science Publ*, 1991.

[11] P. Lelong. Intégration sur un ensemble analytique complexe. *Bull. soc. math. France*, 85 :239–262, 1957.

[12] P. Lelong. D'une variable plusieurs variables en analyse complexe : les fonctions plurisousharmoniques et la positivité. In *Revue d'Histoire des math.*, volume 1, pages 139–157. Soc.Math. de France, Paris, 1995.

[13] F. Norguet. Sur les domaines d'holomorphie des fonctions uniformes de plusieurs variables complexes (passage du local au global). *Bull. Soc. Math. France*, 82 :137–159, 1954.

[14] A. Rashkovskii. Singularities of plurisubharmonic functions and positive closed currents. Reports, N 20 2000, ISSN 1400-4798, 2000.

[15] Y.T. Siu. Analyticity of sets associated to Lelong numbers and the extension of closed positive currents. *Invent. math.*, 27 :53–156, 1974.

[16] H. Skoda. Estimations L^2 pour l'opérateur $\overline{\partial}$ et applications arithmétiques. In *Séminaire Pierre Lelong 1975-1976*, volume 578 of *Lectures Notes in Mathematics*, pages 314–323. Springer-Verlag, 1976.

[17] P. Thie. The Lelong number of a point of a complex analytic set. *Math. Ann.*, 172 :266–312, 1967.

Sur le problème de Plateau complexe

Pierre Dolbeault

1 Introduction

Rappelons que le problème de Plateau classique est la recherche d'une surface de l'espace euclidien $\mathbb{R}^n$, $n \geq 3$, d'aire minima, de bord une courbe fermée, continue, donnée. Il a été complètement résolu, en 1929, par J. Douglas [11] pour une courbe simple, i.e. sans points multiples ; mais il a été à l'origine d'une branche du Calcul des variations toujours active.

Le problème de Plateau complexe est analogue, dans une variété complexe, au problème de Plateau classique. Il s'agit d'un problème très particulier de Géométrie ; outre son intérêt propre, il fait intervenir des Géométries introduites ou développées au cours du 20^e siècle : variétés analytiques complexes, propriétés métriques des sous-variétés réelles de $\mathbb{C}^n$, géométrie kählérienne, théorie géométrique de la mesure, géométrie symplectique et des notions topologiques : $\overline{\partial}$-cohomologie, coefficient d'enlacement.

2 Origine possible : extension holomorphe ; enveloppe polynômiale d'une courbe réelle

1. Les variétés analytiques complexes de dimension 1 ont été définies, de façon tout à fait précise, par Hermann Weyl en 1913. Les variétés de dimension supérieure ont été utilisées par W. V. D. Hodge, dans les années 30, dans le cas des variétés projectives lisses. La définition qui suit n'a été donnée qu'à la fin des années 40.
   Soit X une variété de dimension $2n$, c'est-à-dire un espace topologique séparé dont chaque point possède un voisinage ouvert homémorphe à un ouvert de $\mathbb{R}^{2n}$. Supposons, de plus, qu'il existe un recouvrement de X par des ouverts U_i, $i \in I$ et, $\mathbb{R}^{2n}$ étant identifié à $\mathbb{C}^n$, un homémorphisme f_i de $U_i, (i \in I)$ sur un ouvert de $\mathbb{C}^n$ tel que, pour $U_i \cap U_j \neq \emptyset$, $f_j \circ f_i^{-1}$ soit une application biholomorphe de $f_i(U_i \cap U_j)$ sur $f_j(U_i \cap U_j)$. On dit alors que X est une *variété analytique complexe*, de dimension complexe n, ou encore que c'est une variété de dimension réelle $2n$ munie d'une structure analytique complexe. Un couple (U_i, f_i) est appelé une carte de X et U_i un ouvert de carte.

2. Rappelons la définition de $\overline{\partial}$: dans $\mathbb{C}^n$, ou plus généralement dans un ouvert de carte U_j (section 1.1) d'une variété complexe, sur lequel les coordonnées locales complexes sont $(z_1, \ldots, z_n)$, on a : $\overline{\partial} = \sum_{j=1}^{n} \frac{\partial}{\partial \overline{z}_j} d\overline{z}_j$. On appelle problème du $\overline{\partial}$, ou $\overline{\partial}$-problème, la recherche d'une forme différentielle u satisfaisant à l'équation $\overline{\partial} u = f$ où f est une forme différentielle telle que $\overline{\partial} f = 0$.
Le théorème d'extension de Hartogs, obtenu au début du 20^e siècle, a été complètement démontré par Bochner et Martinelli, indépendamment en 1943. F. Severi semble être le premier à avoir introduit la notion de fonction CR dans l'énoncé (1931). J.-P. Serre (1955), utilisant sa dualité a réduit la preuve à la résolution d'un $\overline{\partial}$-problème, Ehrenpreis (1961) a donné une démonstration par résolution du $\overline{\partial}$ à support compact ; Hörmander (1966) a utilisé la propriété $H^1_c(\mathbb{C}^n, \mathcal{O}) = 0$. La version la plus simple est :
Soit Ω un ouvert borné de $\mathbb{C}^n$, $n \geq 2$. Supposons que $\partial\Omega$ soit une sous-variété connexe de classe C^k $(1 \leq k \leq \infty)$ ou de classe C^ω (i.e. analytique réelle). Soit f une fonction de $C^l(\partial\Omega)$, $1 \leq l \leq k$.
Alors les deux conditions suivantes sont équivalentes :
(i) f est une fonction CR, i.e. la différentielle de f restreinte aux sous-espaces complexes de l'espace tangent à $\partial\Omega$, en tout point, est $\mathbb{C}$-linéaire ;
(ii) il existe $F \in C^l(\Omega) \cap \mathcal{O}(\Omega)$ telle que $F\,|_{\partial\Omega} = f$.
Alors le graphe de f est le bord de la sous-variété analytique complexe définie par le graphe de F dans $\mathbb{C}^{n+1}$.

3. Soit M une sous-variété compacte de dimension 1 de $\mathbb{C}^n$, on appelle enveloppe polynômiale de M, le compact $\{z \in \mathbb{C}^n ;\ \mid P(z) \mid \leq \max_{\zeta \in M} \mid P(\zeta) \mid ; P \in \mathbb{C}[Z]$, l'anneau des polynônes à coefficients complexes$\}$. Alors (J. Wermer (1958)), l'enveloppe polynômiale de M est soit M, soit la réunion de M et du support d'un ensemble analytique complexe T, de dimension complexe 1, de bord M [26].

3 Minimalité des ensembles analytiques complexes dans une variété kählérienne

3.1 Inégalité de Wirtinger (1936)

Dans $\mathbb{C}^n$, on considère les coordonnées complexes $(z_1, \ldots, z_n)$, $z_j = x_j + iy_j$, la métrique hermitienne $H = \sum_{j=1}^{n} dz_j \otimes d\overline{z}_j$ et la forme extérieure (de Kähler standard)

$$\omega = \frac{i}{2} \sum_{j=1}^{n} dz_j \wedge d\overline{z}_j = \sum_{j=1}^{n} dx_j \wedge dy_j.$$

A partir de l'espace vectoriel réel $\mathbb{R}^{2n} \cong \mathbb{C}^n$, on considère l'espace vectoriel réel $\Lambda_{2p}\mathbb{R}^{2n}$ (multivecteurs pour le produit extérieur, avec la norme associée $\mid . \mid$) ; chaque $2p$-vecteur

décomposable définit un $2p$-plan réel de $\mathbb{C}^n$. On définit la norme $||\zeta|| = \inf \sum_j \mid \zeta_j \mid$ où $\zeta = \sum_j \zeta_j$, ζ_j décomposable (i.e. produit extérieur d'éléments de $\mathbb{R}^{2n}$).

Soit $P_{pp} = \{\sum_{j=1}^{N} \lambda_j \zeta_j;\quad \zeta_j$ décomposable définissant un p-plan complexe de $\mathbb{C}^n$; $\lambda_j \geq 0; N \in \mathbb{N}^*\}$.

Alors, pour tout $\zeta \in \Lambda_{2p}\mathbb{C}^n$*, on a :*

$$\frac{1}{p!}\omega^p(\zeta) \leq ||\zeta||;$$

égalité pour $\zeta \in P_{pp}$ *uniquement* [27].

3.2 Variété kählérienne

Une variété kählérienne est une variété analytique complexe X munie d'une métrique hermitienne dont la 2-forme différentielle extérieure associée ω est d-fermée, i.e. $d\omega = 0$, (Kähler, 1933).

L'espace projectif complexe, muni de la métrique de Fubini-Study et les variétés algébriques projectives lisses, munies de la métrique induite, sont des variétés kählériennes.

Rappellons que les *courants* définis par G. de Rham [22], d'abord dans des cas particuliers, dans les années 30, puis, en général, en 1950, à la suite des distributions de L. Schwartz, sont les formes linéaires continues sur l'espace des formes différentielles à support compact sur une variété différentielle riemanniene, muni d'une topologie convenable : sur chaque ouvert de carte, ce sont les formes différentielles à coefficients distribution.

On considère les courants S à coefficients mesure. Leur étude constitue la Théorie géométrique de la mesure [12] décrite dans le gros traité de H. Federer (1969). On définit la masse $M(S)$ d'un tel courant à support compact ; pour tout compact K de X, on définit la masse, sur K, $M_K(S)$ du courant S.

Parmi les espaces de courants considérés figure l'espace des *courants localement rectifiables* obtenus comme limites, pour la topologie définie par la masse, des combinaisons linéaires (localement finies), à coefficients entiers d'images, par des applications de classe C^1, de simplexes de $\mathbb{R}^p$ (triangles de $\mathbb{R}^2$, les tétraèdres de $\mathbb{R}^3$, etc). Le théorème d'intégrabilité sur un ensemble analytique complexe W, de dimension p, de P. Lelong [21] montre que le courant d'intégration $[W] = \int_{Reg W}$ sur W est un courant localement rectifiable, d-fermé, de bidimension (p,p) et positif (voir l'exposé de P. Lelong) . Si S est le courant d'intégration $[V]$ sur un ensemble analytique complexe V, $M_K[V]$ est le volume de $V \cap K$.

Il résulte de l'inégalité de Wirtinger que, pour tout courant d-fermé, à coefficients mesure, à support compact R, on a

$$(*) \qquad M_K[V] \leq M(\chi_K[V] + R)$$

où χ_K est la fonction caractéristique de K ; en outre, le support de $\chi_K[V]$ est $K \cap V$.

L'une des techniques introduites par Federer, utile pour notre problème, est le tranchage que l'on va définir. Soient X, Y deux variétés C^∞, orientées, de dimensions respectives m et n, $m \geq n$, et f une application C^∞ :$X \to Y$; on suppose Y munie de la mesure de Hausdorff n-dimensionnelle. Pour tout courant u, de degré k, $0 \leq k \leq m-n$, localement plat (un peu plus général qu'un courant localement rectifiable), un *tranchage* de u par f est une collection de courants $< u, f, a >$ de X, appelés tranches, définis, pour presque tout $a \in Y$, tels que, pour toute forme différentielle φ, C^∞, à support compact, sur X, pour toute forme C^∞ η de Y, de degré n, l'application : $Y \to \mathbb{C}$, $a \mapsto < u, f, a > (\varphi)$ soit mesurable et intégrable sur Y et satisfasse $\int_Y < u, f, a > (\varphi)\eta = (u \wedge f^*\eta)(\varphi)$, où $f^*\eta$ est l'image réciproque de η par f.

Ex. : si u est (le courant d'intégration défini par) une sous-variété V de X, $< u, f, a > (\varphi) = \int_{f^{-1}(a)\cap V} \varphi$.

4 Problème du bord comme problème d'extension géométrique : problème de Plateau dans le cas kählérien

Chaînes holomorphes Une p-chaîne holomorphe, d'un ouvert Ω d'une variété complexe X, est un courant localement rectifiable, d-fermé, de bidimension (p,p), $T = \sum n_j[W_j]$, où W_j est un ensemble analytique complexe, irréductible, de dimension complexe p, $n_j \in \mathbb{Z}$, la somme étant localement finie. La positivité est vérifiée si les n_j sont ≥ 0.

Remarques

Rappelons qu'une variété à bord (X, bX), de dimension r, est un espace topologique séparé dont chaque point a un voisinage ouvert, dans X, homéomorphe à un ouvert de $\mathbb{R}^r$ ou du demi-espace $\mathbb{R}^r$ fermé égal à $\mathbb{R}^{r-1} \times \mathbb{R}^+$, avec $\mathbb{R}^+ = \{x \in \mathbb{R}, x \geq 0\}$, bX étant envoyé dans le sous-ensemble $\mathbb{R}^{r-1} \times \{0\}$ de $\mathbb{R}^{r-1} \times \mathbb{R}^+$ par les homéomorphismes ci-dessus.

Généralement $(sptT, M)$ n'est pas une variété à bord au voisinage de M, comme le montre un exemple de Harvey et Lawson (1975) décrit ci-dessous [14]. Cette situation oblige à définir le bord au sens des courants.

Soient $\phi : \mathbb{C} \to \mathbb{C}^2$ l'application holomorphe : $t \mapsto (z_1 = t^2, z_2 = t^3 - t)$ et $a = -1$, $b = 1$ dans $\mathbb{C}$.

ϕ est une immersion, $\phi(b) = \phi(a) = (1,0)$, et $\phi_{|\mathbb{C}\setminus\{a,b\}}$ est un plongement. Soit B le disque ouvert de $\mathbb{C}$, de centre a et de rayon 2 ; $M = \phi(\partial B)$ est une sous-variété analytique réelle de $\mathbb{C}^2$. $V = \phi(B \setminus a)$ est une sous-variété analytique complexe de $\mathbb{C}^2 \setminus M$ et $b[V] = M$. Au voisinage de (1,0) dans $\mathbb{C}^2$, l'ensemble des points de V de paramètre t voisin de 1 est une variété de bord M, mais ce n'est pas le cas pour t voisin de -1. (figure 1)

Dans le cas où M n'est pas connexe, en général T est une chaîne holomorphe non définie par un seul ensemble analytique irréductible ; ex. : si M est la réunion de deux cercles concentriques de rayons 1 et 2, suivant leurs orientations, la chaîne holomorphe T a pour composantes ou bien deux disques avec des coefficients 1 et -1 respectivement, ou bien une couronne.

Problème du bord des chaînes holomorphes

Sur une variété analytique complexe X, munie d'une métrique hermitienne, de dimension complexe n, soit M une sous-variété réelle, fermée, orientée, de classe C^k, $k \geq 1$, de dimension $2p-1$; $0 < p \leq n$. On notera encore M le courant d'intégration $[M] = \int_M .$ sur M ; alors M

est de masse localement finie et $dM = 0$. S'il existe une p-chaîne holomorphe T de $X \setminus M$ de masse localement finie au voisinage de M, T admet une extension simple à X, notée encore T et si $dT = M$, on dit que M est le bord de T et que T est une solution du *problème du bord pour M dans X*. En fait, d est égal à $-b$, où b est l'opérateur bord des courants ; quitte à faire une erreur de signe, on utilisera d en général.

M étant donnée sur X, on cherche des conditions nécessaires et suffisantes pour que M soit le bord d'une p-chaîne holomorphe : c'est le *problème du bord pour M dans X*.

Dans la situation de la section 1.2 (graphes des fonctions figurant dans le théorème de Hartogs), la solution du problème de bord apparaît comme une extension de grf en une variété holomorphe (i.e. analytique complexe) traduisant l'extension de f par F. Il en est de même dans le cas général si la chaîne holomorphe a tous ses coefficients égaux à 1.

Problème de Plateau au sens strict dans le cas kählérien

Quand T a tous ses coefficients égaux à 1, d'après la conséquence $(*)$ (section 2.4) de l'inégalité de Wirtinger, sur une variété kählérienne, le volume de l'ensemble analytique *spt* T est minimal ; c'est un minimum relatif, car la solution du problème du bord n'est pas unique en général.

Parfois, on appelle problème de Plateau, le problème du bord pour M dans X.

5 Solutions du problème du bord dans différents espaces

1. Le premier résultat a été obtenu en 1958, par J. Wermer, dans $\mathbb{C}^n$, pour $p = 1$ et M image holomorphe du cercle unité de $\mathbb{C}$ [26] ; ce résultat a été généralisé au cas où M est de classe C^1, réunion de courbes réelles connexes, par Bishop, Stolzenberg (1966), en recherchant l'enveloppe polynomiale de M selon la section 1.3.
   Dans $\mathbb{C}^n$, après des résultats préliminaires de Rothstein (1959) [23], le problème du bord a été résolu par Harvey et Lawson (1975), pour $p \geq 2$, sous la condition nécesaire et suffisante suivante : M est compacte, maximalement complexe et, pour $p = 1$, sous la condition des moments : $\int_M \varphi = 0$, pour toute 1-forme holomorphe φ sur $\mathbb{C}^n$ [15]. Pour $n = p+1$, la méthode, inspirée du théorème de Hartogs-Bochner, consiste à construire T comme diviseur d'une fonction méromorphe elle-même construite, pas à pas, à partir de solutions de $\overline{\partial}$-problèmes à support compact. T peut aussi être considéré comme graphe (avec multiplicités sur les composantes irréductibles) d'une fonction analytique à un nombre fini de déterminations. Pour p quelconque, on se ramène au cas précédent à l'aide de projections.
   Dans $\mathbb{C}P^n \setminus \mathbb{C}P^{n-r}$, $1 \leq r \leq n$, pour M compact, le problème a une solution si et seulement si, pour $p \geq r+1$, M est maximalement complexe et si, pour $p = r$, M satisfait à la condition des moments : $\int_M \varphi = 0$, pour toute $(p, p-1)$-forme φ $\overline{\partial}$-fermée. La méthode consiste à résoudre le problème du bord, dans $\mathbb{C}^{n+1} \setminus \mathbb{C}^{n-r+1}$, pour l'image réciproque de M par la projection canonique [16].
   Dans les deux cas la solution est unique.
   Harvey et Lawson supposent que la donnée M est un compact qui, à l'exception d'un fermé de mesure de Hausdorff $(2p-1)$-dimensionnelle nulle est une variété de classe

C^1, orientée ; on dira que M est une variété C^1 à singularités négligeables.
E. Chirka, en 1985, pour $p \geq 2$ et M relativement compact, a résolu le problème du bord dans le complémentaire d'un compact holomorphiquement convexe dans $\mathbb{C}^n$ [8].
Le problème du bord, dans $\mathbb{C}P^n$ a été posé, pour la première fois par J. King [19] ; l'unicité de la solution n'est plus possible, puisqu'alors, deux solutions diffèrent d'une p-chaîne algébrique.

2. Dans $\mathbb{C}P^n$, une solution au problème du bord a été obtenue par P. Dolbeault et G. Henkin pour $p = 1$, (1994), puis pour p quelconque (1997) et, plus généralement, dans un domaine X q-linéellement concave de $\mathbb{C}P^n$, i.e. réunion de sous-espaces projectifs de dimension q [10]. La condition nécesaire et suffisante d'existence d'une solution est une extension de la condition des moments : elle fait intervenir une formule de résidus de Cauchy (en une variable) et une condition différentielle non linéaire qui apparaît dans de nombreux contextes géométriques ou de Physique mathématique. Dans le cas le plus simple, $p = 1$, $n = 2$, c'est l'équation de l'onde de choc pour une fonction holomorphe locale de 2 variables ξ, η :

$$f\frac{\partial f}{\partial \xi} = \frac{\partial f}{\partial \eta}.$$

A partir d'une condition locale, la relation ci-dessus permet de construire, par extension des coefficients, une fonction méromorphe jouant, dans $\mathbb{C}^n$, le même rôle que la fonction de Harvey-Lawson décrite plus haut ; elle définit une p-chaîne holomorphe qu'on étend à $\mathbb{C}P^n$ à l'aide du théorème classique de Bishop-Stoll.
Pour des raisons techniques, et faute d'utiliser le tranchage, la régularité est supposée C^2.

3. En outre, comme conséquence, et grâce à un théorème de compacité de Sacks-Uhlenbeck, on a : si les sections de M par presque tous les sous-espaces projectifs de codimension $(p+1)$, appartenant à un ouvert assez petit de la Grassmannienne, sont les bords de 1-chaînes holomorphes, alors M est le bord d'une p-chaîne holomorphe de $\mathbb{C}P^n$ ou de X.
Les conditions de régularité de la donnée M, ont été affaiblies, d'abord dans $\mathbb{C}^n$, et pour $p = 1$, en une condition, un peu plus forte que la rectifiabilité, par H. Alexander [2] qui a, en outre, donné un contre-exemple essentiel [1], puis par Lawrence [20] et enfin, et pour p quelconque, dans $\mathbb{C}^n$ et $\mathbb{C}P^n$, par T.C. Dinh [9] : M est un courant rectifiable dont le cône tangent est, presque partout, un sous-espace vectoriel. De plus, Dinh a obtenu la réduction du problème du bord dans $\mathbb{C}P^n$ au cas $p = 1$, avec des conditions moins fortes que ci-dessus et par un procédé analytique élémentaire [9].
Tous les résultats antérieurs sont obtenus comme corollaires.

4. Plus récemment (1999), F. Sarkis a établi une généralisation du résultat 4.2.1 au cas d'un espace $X = U \times \Omega$ où U est une variété complexe connexe, Ω kählérienne compacte, et sous les hypothèses plus larges de Dinh.

Justification de la condition dans $\mathbb{C}P^n$: conditions nécessaires pour $p = 1$

Géométrie de $\mathbb{C}P^n$ Soient $(w_0, w_1, \ldots, w_n)$ des coordonnées homogènes de $\mathbb{C}P^n$. L'hyperplan $Q = \{w_0 = 0\}$ sera dit *hyperplan à l'infini* of $\mathbb{C}P^n$. On considère des coordonnés affines $z_k = \dfrac{w_k}{w_0}$ dans $\mathbb{C}P^n \setminus Q \cong \mathbb{C}^n$. Un hyperplan projectif de $\mathbb{C}P^n$ a l'équation $\tilde{g} = 0$, où

$\tilde{g} = \sum_{l=0}^{n} a_l w_l$; $(a_0, a_1, \ldots, a_n)$ sont les coordonnées homogènes du dual $\mathbb{C}P^{n*}$ de $\mathbb{C}P^n$. On utilisera la carte de domaine $\mathbb{C}P^{n*} \setminus \{a_n = 0\} \cong \mathbb{C}^n$, avec les coordonnées affines

$$\xi = -\frac{a_0}{a_n}, \;\; \eta_l = -\frac{a_l}{a_n}, \;\; l = 1, \ldots, n-1.$$

On pose

$$\eta = (\eta_1, \ldots, \eta_{n-1}), \;\; \nu = (\xi, \eta), \;\; \mu = (w_1, \ldots, w_{n-1}), \;\; \zeta = (z_1, \ldots, z_{n-1}); \;\; \eta.\mu = \sum_{l=1}^{n-1} \eta_l w_l; \;\; \eta.\zeta = \sum_{l=1}^{n-1} \eta_l z_l.$$

Soit D_ν l'hyperplan $\tilde{g}^{-1}(0)$ de $\mathbb{C}P^n$; son équation est $w_n - \xi w_0 - \eta.\mu = 0$. Dans $\mathbb{C}P^n \setminus Q \cong \mathbb{C}^n$, on considère les hyperplans affines $D'_\nu = D_\nu \cap \mathbb{C}^n = g^{-1}(0)$ avec $g = z_n - \xi - \eta.\zeta$.

Lemme de Darboux -*Dans un ouvert $\Omega \subset \mathbb{C}P^n \setminus Q \cong \mathbb{C}^n$, considérons une sous-variété complexe W de dimension 1. Soit ν^* tel que $D'_{\nu^*} \cap W$ soit un ensemble fini. Alors, pour tout ν dans un petit voisinage de ν^*, $D'_\nu \cap W$ est un nombre fini de points $(f_j(\nu), \xi + \eta.f_j(\nu))$, $j = 1, \ldots, N$, où $f_j =^t (f_{jk})$, $k = 1, \ldots, n-1$, est holomorphe en ν et satisfait*

$$(i) \qquad f_{jk}\frac{\partial f_j}{\partial \xi} = \frac{\partial f_j}{\partial \eta_k}; \qquad k = 1, \ldots, n-1$$

Preuve.- Soit $H^{-1}(0) = W$ où $H(\zeta; z_n)$ est holomorphe dans Ω ; alors $\mathcal{H}(\nu) = H(f_j(\nu), \xi + \eta.f_j(\nu)) \equiv 0$; les conditions de compatibilité pour vérifier $d\mathcal{H} \equiv 0$ donnent (i). □

Pour $n = 1$, (i) est l'equation de l'onde de choc.

Ce Lemme est implicite dans Darboux (Théorie des surfaces I).

Formule de résidus de Cauchy

Soit $M = \gamma$ une courbe réelle C^1 de volume 1-dimensionnel fini. On choisit Q tel que $\gamma \cap Q = \emptyset$. On suppose que $\gamma = d(1.[S])$ où S une courbe analytique complexe. On pose aussi $\gamma = dS$.

Presque tout $D_\nu \cap S$ est un ensemble fini de points que l'on énumère arbitrairement : $p_j = (f_j(\nu), \xi + \eta.f_j(\nu))$, $j = 1, \ldots, N$ sont définis comme ci-dessus avec S à la place de W. Pour presque tout Q, il existe un nombre fini de points d'intersection q_s, $s = 1, \ldots, N^-$ de Q et de S. On pose

$$G(\nu) = \frac{1}{2\pi i}\int_\gamma \zeta \frac{dg}{g}.$$

Soit ω_1 la forme $\zeta\dfrac{dg}{g}$ en coordonnées homogènes ;

Pour presque tout ν, D_ν est transverse à S et à distance finie, $\omega_{1|S}$ a les pôles simples p_j et, à l'infini, $\omega_{1|S}$ a les pôles q_s. Sur S, considérons les disques B_j, B'_s de rayon ε centrés en p_j et q_s respectivement. Soit B la réunion de ces disques. Alors $b(S \setminus B) = bS - bB$ et $\omega_1 \mid_{S\setminus B}$ est C^∞ et fermée. Un calcul classique de résidus donne : $\lim\limits_{\varepsilon \to 0} \int_{bB_j} \omega_1 = 2\pi i f_j(\nu)$

et $\lim_{\varepsilon \to 0} \int_{bB'_s} \omega_1$ est linéaire en ξ, donc sa dérivée par rapport à ξ est 0. D'où la relation

$$\frac{\partial^2}{\partial \xi^2} G(\nu) = \frac{\partial^2}{\partial \xi^2} \sum_{j=1}^{N} f_j(\nu)$$

□ (figure 2)

Une condition topologique dans $\mathbb{C}^n$

Récemment (2000), H. Alexander et J. Wermer [3] ont établi une version du théorème du bord dans $\mathbb{C}^n$ sans condition analytique apparente.

Coefficient d'enlacement : (linking number)

Soient M une sous-variété de dimension impaire k, orientée, compacte de $\mathbb{C}^n$ et A une sous-variété algébrique de $\mathbb{C}^n$ de dimension complexe s, telle que $k+2s = 2n-1$, munie de son orientation naturelle. A étant non compacte, on la remplace par une variété compacte A' coïncidant avec A à l'intérieur d'une boule de rayon assez grand. Soit Σ une $(s+1)$-sous-variété orientée, compacte de $\mathbb{R}^{2n} \cong \mathbb{C}^n$ telle que $M = b\Sigma$, le coefficient d'enlacement de M et de A est

$$link(M, A) = \sharp(\Sigma, A'),$$

nombre d'intersection de Σ et de A' (figure 3).

Théorème - *Soit M comme ci-dessus, telle que $3 \leq k \leq 2n-3$. Alors les conditions suivantes sont équivalentes :*

(i) $$link(M, A) \geq 0$$

pour toutes les sous-variétés algébriques de $\mathbb{C}^n$ disjointes de M, de dimension pure complexe : $n - \frac{k+1}{2}$;

(ii) M est maximalement complexe et (par Harvey-Lawson) *il existe une unique $\frac{k+1}{2}$-chaîne holomorphe, dans $\mathbb{C}^n \setminus M$, de masse finie, de support borné telle que $M = bT$.*

De plus, pour tout $x \in \mathbb{C}^n \setminus M$, $x \in spt\ T$ équivaut à

$$link(M, A) > 0$$

pour toute A comme ci-desus, telle que $x \in A$.

Démonstration

En utilisant Harvey-Lawson, d'abord, pour $k = 1$, en faisant intervenir la condition du moment, puis pour $k = 3$, enfin pour le cas général par tranchage. Un lemme essentiel, pour $k = 1$, $M = \gamma$ dit :

$$link(\gamma, A) = \frac{1}{2\pi i} \int_{\gamma} \frac{dP}{P}$$

où $P \in \mathbb{C}[Z]$ et $A = Z(P)$.

Problème du bord des hypersurfaces Levi-plates dans $\mathbb{C}^p$: cas de $\mathbb{C}^2$; un exemple trivial et le feuilletage de Jöricke dans $\mathbb{C}^3$

On s'intéresse au problème général suivant : soit M une surface dans $\mathbb{C}^2 = \{(z, w), z = x+iy, w = u+iv\}$, trouver une hypersurface X Levi-plate (i.e. feuilletée par des courbes complexes) dans $\mathbb{C}^2 \setminus M$, telle que $bX = M$. La méthode des disques analytiques de Bishop (i.e.

images holomorphes du disque unité de $\mathbb{C}$) [7] joue un rôle essentiel dans les démonstrations d'une partie des résultats.

Etant donnée une partie E de $\mathbb{C}^2$, soit $\mathcal{O}(E)$ l'ensemble des germes de fonctions holomorphes définies au voisinage de E. On appelle *enveloppe d'holomorphie* le spectre (i.e. l'espace des idéaux maximaux) de $\mathcal{O}(E)$.

Lorsque M est lisse et compacte, le premier résultat obtenu (Bedford-Gaveau, 1983) [4] est le suivant :

(a) Soient G un domaine borné (à bord) lisse de $\mathbb{C}_z \times \mathbb{R}_u$, M le graphe $\Gamma(\varphi)$ d'une fonction lisse φ sur ∂G tels que :

(b) M est une 2-sphère topologique avec exactement 2 points elliptiques où le plan tangent est complexe (en un tel point, pour des coordonnées complexes (z, w) convenables, M a pour équation $w = z\overline{z} + \beta \mathcal{R}ez^2 + O(\mid z \mid^3)$ où $0 \leq \beta < 1$) ;

(c) $\partial G \times \mathbb{R}_v$ est strictement pseudoconvexe (définition dans l'exposé de P. Lelong).

Alors, $\Gamma(\varphi)$ est le bord d'un graphe Levi-plat $\Gamma(\Phi)$ où Φ est une fonction lisse dans G et Lipschitz sur $\overline{G}$. En outre $\Gamma(\Phi)$ est engendré par une famille de disques analytiques et $\overline{\Gamma(\Phi)}$ est l'enveloppe d'holomorphie de $\Gamma(\varphi)$. (figure 4)

Il y eut ensuite la généralisation suivante : M n'est pas nécessairement un graphe mais une 2-sphère topologique contenue (condition nécesaire) dans le bord d'un domaine strictement pseudoconvexe. Alors, à une perturbation près, M borde une hypersurface Levi-plate (Bedford-Klingenberg, Eliashberg et Kruzilin, indépendamment, en 1991, Gardiner en 1993). Une présentation de ces résultats se trouve dans [6].

Supposant φ seulement continue, G borné, strictement pseudoconvexe dans $\mathbb{C}_z \times \mathbb{R}_u$, Shcherbina (1993) a résolu le problème initial avec Φ continue et des disques analytiques, puis Chirka-Shcherbina (1995) ont généralisé ce résultat.

En 1991, dans l'hypothèse : $G \times \mathbb{R}_v$ strictement pseudoconvexe, $\Phi : \overline{G} \to \mathbb{R}_v$ continue et $\Phi \mid_{\partial G} = \varphi$, H. Alexander et Slodkowski-Tomassini ont montré que $\Gamma(\Phi)$ est l'enveloppe polynomiale de $\Gamma(\varphi)$ si et seulement s'il divise $G \times \mathbb{R}_v$ en deux domaines pseudoconvexes.

Un excellent historique de ce problème se trouve dans [25].

Très récemment (1997), Shcherbina et Tomassini ont étudié le problème général dans le cas $M = \Gamma(\varphi)$ où φ est une fonction continue sur le bord d'un domaine non borné G de $\mathbb{C}_z \times \mathbb{R}_u$ en supposant $G \times \mathbb{R}_v \subset \mathbb{C}^2_{z,w}$ strictement pseudoconvexe, alors le problème a des solutions $X = \Gamma(\Phi)$ où Φ est continue mais en général n'est pas unique.

Les graphes de Levi donnent des solutions faibles pour l'opérateur de Levi L et le résultat obtenu fournit, en particulier, un théorème d'existence pour le problème de Dirichlet pour L [24].

Deux exemples dans $\mathbb{C}^3$

Il n'y a pas, actuellement de resultats généraux dans $\mathbb{C}^3$. Voici deux exemples de situations très différentes.

1. Dans $\mathbb{C}^3$, soit

$$(M) \qquad \begin{cases} z_1\overline{z}_1 + z_2\overline{z}_2 + z_3\overline{z}_3 &= 1 \\ z_3 &= \overline{z}_3 \end{cases}$$

On a CR-*dim* $M = 1$ sauf aux points $z_1 = z_2 = 0; z_3 = \pm 1$ où CR-*dim* $M = 2$. M est la sphère unité de $\mathbb{C}^2 \times \mathbb{R}$; elle borde la boule unité T de $\mathbb{C}^2 \times \mathbb{R}$, qui est feuilleté par les boules complexes $\mathbb{C}^2 \times \{x_3\} \cap T$. Les feuilles sont relativement compactes de dimension réelle 4 et sont bordées par les feuilles compaces (3-sphères) d'un feuilletage de M.

2. Jöricke décrit un exemple élaboré ([J 99], section 5) ayant, en particulier, les propriétés suivantes :
   Il existe un domaine strictement pseudoconvexe Ω dans $\mathbb{C}^3$ de frontière $\partial\Omega$ de classe C^∞ et une sous-variété propre, CR générique, C^∞, compacte, connexe M de $\partial\Omega$, $dim_{\mathbb{R}}\ M = 4$, telle que :
   1) M possède un feuilletage lisse de codim. 1, à feuilles sous-variétés CR de M ;
   2) il n'y a pas de feuille compacte ;
   3) $M = \mathcal{N} \cap \partial\Omega$, où $\mathcal{N}$ est une 5-sous-variété d'un voisinage de $\overline{\Omega}$, munie d'un feuilletage de codim. 1 induisant celui de M, sans feuille relativement fermée, chaque feuille étant, localement, une variété complexe de dimension complexe 2.

6 Problème du bord et géométrie symplectique

Variété symplectique [13]

Soit V une variété lisse, de dimension (réelle) $2n$, munie d'une structure *presque complexe* J, c'est-à-dire d'un champ C^∞ de strucures complexes sur les espaces tangents $T_v(V), v \in V$.

Une 2-forme différentielle extérieure ω sur une variété presque complexe (V, J) est dite *J-positive* si, pour tout point $v \in V$, ω est strictement positive sur toute droite complexe de $T_v(V) \cong \mathbb{C}^n$, i.e. $\omega(\xi, J\xi) > 0$ pour tout $0 \neq \xi \in T_v(V)$.

Une 2-forme différentielle ω, d-fermée sur V est dite *symplectique* si ω^n esr partout non nulle ; (V, ω) est alors appelée une *variété symplectique*

On appelle *domaine de Grauert* ouvert (resp. fermé) un ouvert strictement pseudoconvexe, relativement compact d'une variété analytique complexe (resp. son adhérence). Un tel domaine s'obtient toujours par éclatement d'un nombre fini de points d'un espace de Stein.

Théorème - *Soient Ω un domaine de Grauert de dimension complexe 2, de bord $\partial\Omega$ et Σ une sphère de dimension 2 plongée dans $\partial\Omega$. Alors, il existe une petite déformation C^2 (i.e. 2 fois continûment différentiable) Σ' de Σ et une boule B', de dimension 3, réunion disjointe de disques analytiques complexes (i.e. images holomorphes du disque unité de $\mathbb{C}$) telles que $bB' = \Sigma'$.*

Si Ω est de Stein, $B' \cup \Sigma'$ est l'enveloppe d'holomorphie de Σ' par rapport à $\Omega \cup \partial\Omega$.

Cet énoncé [6] est intermédiaire entre le théorème analytique complexe de Bedford-Klingenberg [5] et un théorème valide pour des domaines presque complexes positifs (Eliashberg), d'après le lemme suivant : *En dimension 2, un domaine de Grauert fermé G est positif (i.e. il existe une forme symplectique ω sur G strictement positive par rapport à la structure complexe J.*

La preuve du théorème est une analyse des disques analytiques de bord sur Σ', à la suite de nombreux travaux, en particulier de Bishop, Bedford, Gaveau, Gromov.

Bibliographie

[1] H. Alexander, *Polynomial hull and linear measure*, Lecture Notes in Math., **1276**, Springer Verlag, New-York (1987), 1-11.

[2] H. Alexander, *The polynomial hull of a rectifiable curve in* $\mathbb{C}^n$ Amer. J. Math., **110** (1988), 629-640.

[3] H. Alexander and J. Wermer,*Linking numbers and boundaries of varieties*. Ann. of Math. (2) **151** (2000), 125-150.

[4] E. Bedford and B. Gaveau,*Envelopes of holomorphy of certain 2-spheres in* $\mathbb{C}^2$, Am. J. Math. **105** (1983), 975-1009.

[5] E. Bedford, W. Klingenberg, *On the envelope of holomorphy of a 2-sphere in* $\mathbb{C}^2$. J. Amer. Math. Soc. **4** (1991), 623–646.

[6] D. Bennequin, *Topologie symplectique, convexité holomorphe et structures de contact d'après Y.* Eliashberg, D. McDuff et al., Sém. Bourbaki 1989-90, *Astérisque* **189-190**, exposé 725 (1990), 285-323.

[7] E. Bishop, *Differentiable manifolds in complex Euclidean space*, Duke Math. J. **32** (1965),1-22.

[8] E.M. Chirka, *Complex analytic sets*, Mathematics and its applications, **46**, Kluwer Academic Publishers, 1989 ; Rusian edition 1985.

[9] T.C. Dinh, *Enveloppe polynomiale d'un compact de longueur finie et chaînes holomorphes à bord rectifiable*, Acta Math. **180** (1998), 31-67.

[10] P. Dolbeault et G. Henkin,*Chaînes holomorphes de bord donné dans* $\mathbb{C}P^n$, Bull. Soc. Math. France **125** (1997), 383-445.

[11] J. Douglas, *Solution of the problem of Plateau*, Trans. Am. Math. Soc. **33** (1931), 263-321.

[12] F. Federer, *Geometric measure theory*, Grundlehren der Math. Wiss., **285**, Sringer Verlag, 1969.

[13] M. Gromov, *Pseudo-holomorphic curves in symplectic manifolds*, Invent. Math. **82** (1985), 307-347.

[14] R. Harvey,*Holomorphic chains and their boundaries*, Proc.Symp. Pure Math., **30**, part 1, Am. Math. Soc., Providence R.I. (1977), 309-382.

[15] R. Harvey and B. Lawson, *On boundaries of complex analytic varieties*, I, Ann. of Math., **102**, (1975), 233-290.

[16] R. Harvey and B. Lawson, *On boundaries of complex analytic varieties*, II, Ann. of Math., **106**,

[17] B. Jöricke, *Boundaries of singularity sets, removable singularities, and CR-invariant subsets of CR- manifolds*, J. Geom. Anal. **9** (1999), no. 2, 257–300.

[18] E. Kähler, *Über eine bemerkenswerte Hermitische Metrik*, Abh. Math. Sem. Univ. Hamburg **9** (1933), 173-186.

[19] J. King, *Open problems in geometric function theory*, Proceedings of the fifth international symposium of Math.,p. 4, The Taniguchi foundation, 1978.

[20] M.G. Lawrence,*Polynomial hulls of rectifiable curves*, Amer. J. Math., **117** (1995), 405-417.

[21] P. Lelong, *Intégration sur un ensemble analytique complexe*, Bull. Soc. Math. France, **85** (1957), 239-262.

[22] G. de Rham, *Variétés différentiables*, Paris Hermann, 1955.

[23] W. Rothstein, *Bemerkungen zur Theorie komplexer Räume*, Math. Ann., **137** (1959), 304-315.

[24] N.V. Shcherbina and G. Tomassini, *The Dirichlet problem for Levi-flat graphs over unbounded domains*. Internat. Math. Res. Notices 1999, no. 3, 111–151.

[25] G. Tomassini,*Boundaries of Levi-flat hypersurfaces of* $\mathbb{C}^2$. Complex analysis and geometry, 213–219, (Paris, 1997), Progress in Mathematics, **188**, Birkhäuser Verlag, Basel, 2000.

[26] J. Wermer, *The hull of a curve in* $\mathbb{C}^n$. Ann.of Math. **68** (1958), 550-561.

[27] W. Wirtinger,*Ein Determinantenidentität und ihre Anwendung auf analytische Gebilde und Hermitische Massbestimmung*. Monatsh. f. Math. u. Physik, **44** (1936), 343-365.

L'influence de la théorie de Galois sur l'œuvre de Grothendieck

François CHARGOIS

A la mémoire de Chantal

1 Introduction

Le présent texte est la rédaction d'une conférence prononcée à l'occasion du colloque international "Géométrie au vingtième siècle : 1930-2000", qui s'est tenu à Paris en septembre 2001. Il se propose d'évoquer quelques-unes des influences qu'a exercées la théorie de Galois sur les travaux de Grothendieck.

Si on désire en effet appréhender l'œuvre considérable d'un Titan des Mathématiques tel que Grothendieck, si on ambitionne de parvenir à s'y orienter et d'y tracer son chemin, il n'est certes pas inutile de tenter de raccrocher cette œuvre à quelques grandes thématiques, quelques puissantes idées-force, comme dirait le Maître. Parmi les nombreuses possibilités, parmi les multiples pistes qui s'offrent naturellement, parmi les diverses grilles que l'on peut songer à employer pour décrypter ses travaux, l'œuvre de son illustre devancier Galois est l'une des premières et des plus incontestables qui se présentent à l'esprit.

Un des aspects les plus frappants et intéressants en effet dans l'étude de la création grothendieckienne, est le caractère manifeste de cette influence de Galois sur celle-ci, d'autant plus manifeste que l'on pénètre plus profondément dans cette création, d'autant plus évidente qu'on l'accompagne dans son évolution chronologique. On peut même, croyons-nous, parler d'une véritable parenté de pensée. Elle est totalement consciente et même revendiquée par Grothendieck, comme le montrent certains passages de "Récoltes et Semailles". Nombre de théories de Grothendieck sont donc modelées sur la théorie de Galois classique, ou plutôt sur la reformulation de celle-ci trouvée par Grothendieck dans les années cinquante. C'est dire qu'un rôle-clef y est joué par des groupes de Galois, ou des objets étroitement apparentés, groupoïdes, schémas en groupes, gerbes...

Toutes ces théories, et leur nombreuses variantes, forment un vaste ensemble, que l'on peut baptiser théorie de Galois-Grothendieck, encore que cette dénomination soit ambiguë. Celle-ci constitue donc une extension et un approfondissement considérables de la théorie initiale de Galois. Sous ce rapport comme sous beaucoup d'autres, Grothendieck est allé plus loin que ses prédécesseurs et est parvenu, nous semble-t'il, au cœur de l'essence mathématique ; il nous a livré les clefs d'une véritable compréhension de la pensée de Galois, comme, en une autre occasion, il a révélé à la communauté des mathématiciens celles de l'œuvre de Riemann, un visionnaire de génie lui aussi.

Dans la suite de ce texte, nous nous proposons donc d'exposer quelques-uns des aspects galoisiens de l'œuvre de Grothendieck. Nous expliquerons pour commencer sa reformulation de la théorie de Galois et de la théorie topologique des revêtements et du groupe fondamental de Poincaré . Nous expliquerons ensuite comment, passant des corps aux schémas généraux, Grothendieck a pu formuler une théorie du groupe fondamental en Géométrie Algébrique, première étape d'une théorie de l'homotopie des schémas. Cette théorie du groupe fondamental contient la théorie de Galois initiale et est elle-même contenue dans une théorie axiomatique des catégories galoisiennes, que nous présentons également. L'étape suivante nous amène au cœur de la création grothendieckienne, la théorie des motifs, centrée sur les notions de catégorie tannakienne, de gerbe et de groupe de Galois motivique, que nous esquissons. Pour terminer nous dirons quelques mots d'un programme de recherche plus récent conçu par Grothendieck dans les années soixante-dix et quatre-vingt, nouvelles variations sur le thème galoisien : géométrie algébrique anabélienne, théorie de Galois-Teichmüller, dessins d'enfants...

Il est parfaitement clair qu'il est vain d'espérer faire le tour, dans ce bref article, d'une œuvre aussi considérable, aussi bien quantitativement que qualitativement, que celle de Grothendieck, ou même d'espérer épuiser le sujet proposé ici, celui des aspects galoisiens de cette œuvre : sans doute eût-il fallu par exemple parler également des topos. D'ailleurs, dans le jeu des parentés intellectuelles, nous aurions pu nous tourner également, outre Galois, vers d'autres grands noms des Mathématiques, ne serait-ce que Riemann ou Poincaré , déja cités. Le présent texte ne doit donc être considéré que comme une modeste ébauche. Par ailleurs, il faut certainement regretter que dans un tel colloque, dédié à la géométrie du vingtième siècle, ne se soit trouvée finalement programmée qu'une seule conférence consacrée à Grothendieck. D'autres mathématiciens, infiniment plus qualifiés que l'auteur, auraient pu avec avantage, et sans doute dû, s'y exprimer. Qu'il en ait été autrement est la seule raison de la présence de notre modeste contribution en aussi prestigieuse compagnie...

Précisons également, si tant est que cela ne soit pas évident, que l'auteur ne prétend nullement, en la circonstance, faire œuvre historique ou philosophique, n'étant en rien qualifié en ces matières. En particulier, l'ordre dans lequel sont esquissées ci-dessous quelques-unes des idées de Grothendieck est un ordre thématique, "logique", il exprime parfois une filiation, mais ne prétend pas être un ordre historique.

Signalons que nous n'avons pas su résoudre dans ces notes le problème le plus épineux de la théorie de Galois-Grothendieck : quel est le pluriel du mot "foncteur-fibre" ?

2 La théorie de Galois classique et la reformulation de Grothendieck

2.1 Rappels sur la théorie de Galois

La théorie de Galois classique, telle que l'ont reformulée les successeurs de Galois du dix-neuvième siècle et de la première moitié du vingtième siècle, étudie les extensions (finies séparables) d'un corps commutatif et la manière dont elles s'imbriquent les unes dans les autres. Le mérite de Galois est, en substance, d'avoir compris que cette structure se trouve parfaitement reflétée dans celle d'un groupe qu'il associe au corps considéré, ce qui permet immédiatement l'application des puissants moyens de la théorie des groupes — dont ces travaux de Galois marquent également l'avènement — à l'algèbre et à l'arithmétique. L'impact sur l'arithmétique en particulier est considérable et conduit à la théorie des corps de classes, l'une des cimes des mathématiques de l'époque (et même de la nôtre).

Plus précisément, fixons un corps (commutatif) k, une clôture séparable k^s de k et formons le groupe G des k-automorphismes de k^s : c'est "le" groupe de Galois (absolu) de k. Il se trouve muni naturellement d'une structure de groupe profini, i.e. de groupe topologique compact totalement discontinu. La théorie fournit alors une bijection décroissante naturelle (la correspondance de Galois) de l'ensemble ordonné des extensions finies de k dans k^s sur celui des sous-groupes ouverts de G et plus généralement, avec Krull, de l'ensemble ordonné des extensions de k dans k^s sur celui des sous-groupes fermés de G ; de plus, dans cette bijection, les extensions galoisiennes correspondent aux sous-groupes normaux. Tel est le résultat central de la théorie.

2.2 Analogies avec la théorie topologique des revêtements

A l'orée des années cinquante, lorsque Grothendieck débute sa carrière de mathématicien, les analogies entre la théorie algébrique précédente et la théorie topologique des revêtements et du groupe fondamental, initiée par Poincaré , sont déjà bien connues, mais sans doute peu comprises. Les deux théories se sont de plus trouvées déjà intimement liées dans la théorie des fonctions algébriques de Riemann et ses successeurs. Elles restent néanmoins séparées et c'est Grothendieck qui va accomplir la synthèse.

La théorie topologique cherche à classifier les revêtements d'un espace topologique (raisonnable) connexe. A un tel espace, on associe, une fois choisi un point de celui-ci, un groupe, le groupe fondamental, ou groupe de Poincaré , de cet espace pointé. Les (classes d'isomorphismes de) revêtements connexes de l'espace considéré correspondent alors bijectivement aux (classes de conjugaison de) sous-groupes du groupe de Poincaré et, plus généralement, les (classes d'isomorphismes de) revêtements correspondent bijectivement aux (classes d'isomorphismes d') opérations du groupe de Poincaré sur un ensemble. Il semble d'ailleurs que ce dernier point reste dans une large mesure peu familier aux topologues de l'époque, qui ont tendance à se limiter aux revêtements connexes. Il y a une bonne raison à cela et il semble bien que la clarté ne viendra qu'après l'intervention de Grothendieck. Par ailleurs, la théorie géométrique générale des fibrés et l'opération fondamentale de celle-ci, la construction du fibré associé à un fibré principal et à une action de son groupe structural, qui relève du même cercle d'idées, sont balbutiantes à cette époque et le lien avec Galois certainement

peu visible.

L'analogie entre théorie de Galois et théorie de Poincaré ainsi formulées est en revanche parfaitement explicite : aux corps correspondent les espaces connexes, au choix d'une clôture séparable celui d'un point-base, aux extensions finies séparables les revêtements connexes, au groupe de Galois le groupe de Poincaré et ainsi de suite. Reste à comprendre et à dépasser cette analogie en repensant les deux théories et en les subordonnant à une même théorie plus générale. C'est au passage ce que va accomplir Grothendieck.

2.3 La théorie de Galois revue par Grothendieck

En reformulant les idées de Galois et de Poincaré , Grothendieck va les unifier, en faire deux illustrations d'une même théorie générale, celle des catégories galoisiennes. En réalité et plus exactement, cette théorie générale existe sous plusieurs formes ; la variante qui nous intéresse ici est la variante profinie ; elle englobe effectivement la théorie de Galois dans sa version grothendieckienne et, plus généralement, son extension aux schémas, la théorie du groupe fondamental étale ; la théorie topologique est quant à elle englobée par la variante discrète.

En quoi donc consiste la reformulation de Grothendieck ?

Pour celui-ci, sans doute inspiré par la géométrie, il n'est pas naturel de se limiter aux **extensions** (finies séparables) du corps k considéré, de même qu'il n'est pas raisonnable de se limiter aux revêtements **connexes** d'un espace topologique. Il faut admettre dans la théorie toutes les **k-algèbres** finies séparables, k-algèbres séparables qu'il redéfinit comme étant les k-algèbres universellement réduites (i.e. les k-algèbres A telles que la k-algèbre $A \otimes_k K$ soit réduite pour toute extension K de k). Celles qui sont finies sont exactement les k-algèbres A qui deviennent diagonales par passage à une extension algébriquement close Ω de k, i.e. de la forme Ω^S, où S est un ensemble fini, le **spectre** de A. Elles coïncident également avec les k-algèbres **étales**, la notion générale de morphisme étale étant une des notions fondamentales que Grothendieck introduit à cette époque dans sa théorie des schémas ; nous la retrouverons ci-dessous. Plus concrètement, ces algèbres s'avèrent n'être autres que les produits finis d'extensions finies séparables, comme de juste.

De plus, et corrélativement, ce passage des extensions aux algèbres (et des revêtements connexes aux revêtements généraux) doit s'accompagner de l'abandon du point de vue trop étriqué des ensembles ordonnés au profit de celui des catégories, beaucoup plus souple et plus riche. C'est là une des grandes nouveautés de la contribution de Grothendieck et un pas **décisif**, techniquement très simple, mais d'une portée considérable. Là comme ailleurs, il est devenu indispensable d'introduire, à côté d'objets plus ou moins classiques, la notion appropriée de morphisme. Autrement dit, cette reformulation représente un glissement subtil, qui nécessite de penser non plus en termes d'ensembles mais en termes de catégories, à moins que ce soit **parce que** Grothendieck pense en termes de catégories et non pas en termes d'ensembles, qu'il est amené à cette transition.

Fixons une extension algébriquement close Ω du corps k. L'idée de Grothendieck est alors d'associer à toute k-algèbre étale A son **spectre**, ou plutôt celui de $A \otimes_k \Omega$, qui n'est autre que l'ensemble fini

$$\omega(A) := Hom_{k-alg}(A, \Omega)$$

des caractères de $A \otimes_k \Omega$ (noter l'analogie avec la théorie de Gelfand, bien entendu non fortuite). On obtient ainsi un **foncteur** contravariant ω de la **catégorie** Et_k des k-algèbres étales vers la catégorie Fin des ensembles finis, le foncteur-fibre défini par Ω. Les automorphismes de ce foncteur forment un groupe profini G. Pour Grothendieck, c'est ce groupe qui mérite le nom de groupe de Galois de k (défini par Ω). Il est canoniquement isomorphe au groupe de Galois "ordinaire" $Gal(k^s/k)$, où k^s est la clôture séparable de k dans Ω. Il opère continûment par construction sur les "fibres" $\omega(A)$. Le théorème fondamental de la théorie, qui raffine et enrichit celui de l'ancienne, est alors l'énoncé suivant : ω induit une anti-équivalence de Et_k sur la catégorie Fin_G des ensembles finis E munis d'une action continue (pour la topologie discrète de E) de G.

2.4 Bref retour sur la théorie topologique

La théorie des revêtements et du groupe de Poincaré se reformule alors de façon totalement analogue : pour un espace connexe S comme ci-dessus, le rôle du point-base s est de définir un "foncteur-fibre" ω de la catégorie Rev_S des revêtements de S vers la catégorie Ens des ensembles, à savoir celui qui, à un revêtement X de S, associe sa fibre X_s au point s ; une "meilleure" définition du groupe fondamental $G = \pi_1(S, s)$ se présente alors : c'est le groupe des automorphismes du foncteur-fibre ω. Finalement, on obtient le résultat central de la théorie, qui classifie les revêtements de S : ω induit une équivalence de Rev_S sur la catégorie Ens_G des ensembles munis d'une opération de G.

Remarquons que, pour fabriquer un foncteur-fibre, on peut en fait utiliser un espace T au-dessus de S connexe et simplement connexe arbitraire, pas nécessairement réduit à un point, un revêtement universel par exemple, de même qu'on peut utiliser une extension algébriquement close arbitraire en théorie de Galois. Cette liberté a son utilité. La théorie de la fibration de Milnor en offre une illustration frappante : il y a tout intérêt pour la clarté à y utiliser, plutôt qu'un point-base dans le disque épointé D^* du plan complexe, le demi-plan supérieur disons,

$$\mathcal{H} \longrightarrow D^* \ , \ z \longmapsto s = e^{2\pi i z}$$

qui en est un revêtement universel, ce qui est tout de même plus canonique.

2.5 Quelques conséquences des idées de Grothendieck

La reformulation par Grothendieck de la théorie de Galois, telle que je viens de l'esquisser, a pour vertu première de la rendre fonctorielle et de lui conférer un statut et une saveur incontestablement géométriques. Elle va donc désormais, elle qui relevait de l'algèbre, être plaçée sous l'égide de la géométrie et acquérir ainsi un pouvoir suggestif remarquable. Cette révision de la théorie en fait un élément de la vaste synthèse de l'algèbre, de la topologie, de la géométrie et de l'arithmétique que Grothendieck accomplit à partir de 1955 environ. Les temps sont mûrs pour l'avènement de la Géométrie Arithmétique ; celle-ci contient pour commencer, une "géométrie galoisienne", épanouissement naturel de la théorie de Galois : cohomologie galoisienne, créée et puissamment développée par Tate et Serre vers la même époque, notions appropriées de fibré, de classe caractéristique, etc. De plus, sous cette nouvelle forme, la théorie est d'une certaine façon plus proche de Galois lui-même qu'elle ne l'était devenue aux mains de ses successeurs. En effet, l'intérêt de Galois va aux équations

algébriques séparables et pas seulement à celles qui sont (de plus) irréductibles. Si P est un polynôme séparable à coefficients dans k, l'algèbre $A := k[X]/(P)$ est étale et son spectre $\omega(A)$ n'est autre que l'ensemble R des racines de P dans Ω. La théorie fournit alors une action continue du groupe de Galois G sur R. Son image, un sous-groupe du groupe des permutations de R, est exactement le groupe de Galois de P relativement à k, au sens où Galois lui-même l'entendait. C'est ainsi que les idées de Grothendieck illuminent même la version la plus classique de la théorie de Galois.

Enfin la nouvelle formulation a le mérite d'attirer l'attention de Grothendieck sur les catégories de la forme Fin_G (catégorie des G-espaces discrets finis, voir ci-dessus) pour un groupe profini G (et leurs variantes). Est-il possible de les caractériser intrinsèquement ? La réponse, positive, est l'objet de l'étape suivante, la théorie des catégories galoisiennes, que nous abordons maintenant.

3 Catégories galoisiennes et groupe fondamental étale des schémas

3.1 Catégories et dualité galoisiennes

Comme nous venons de le dire, c'est en analysant les catégories de la forme Fin_G que Grothendieck parvient au concept essentiel de catégorie galoisienne. Une catégorie C est dite **galoisienne** si, outre une propriété de factorisation trop technique pour être détaillée ici, elle jouit des deux propriétés suivantes :

(1) C est exacte, i.e. les limites (projectives et inductives) finies sont représentables dans C,

(2) C admet un foncteur-fibre, i.e. un foncteur $\omega : C \longrightarrow Fin$ tel que :

(i) ω est exact, i.e. commute aux limites (projectives et inductives) finies,

(ii) ω est conservatif, i.e. pour tout morphisme u de C, u est inversible si $\omega(u)$ l'est.

Vérifier que les catégories Fin_G sont galoisiennes et que le foncteur d'oubli $\omega : Fin_G \longrightarrow Fin$ est un foncteur-fibre est facile. Le résultat central de la théorie est la réciproque : les catégories galoisiennes sont exactement les catégories équivalentes à une catégorie de la forme Fin_G pour un groupe profini G convenable. Plus précisément, si C est une catégorie galoisienne, pour tout choix d'un foncteur-fibre ω, on obtient un groupe G, le groupe des automorphismes de ce foncteur ; celui-ci possède une structure naturelle de groupe profini, d'où une catégorie Fin_G ; enfin ω induit une équivalence de C sur Fin_G.

Un point important est la non-unicité des foncteurs-fibres ; ils sont en fait légion. Mais d'autre part, deux quelconques tels foncteurs sont toujours isomorphes, quoique **non canoniquement**. Ces "difficultés" sont déjà présentes dans les situations classiques de Galois et Poincaré mais n'avaient pas été correctement analysées. La manière dont Grothendieck les résoud (et en fait un avantage plutôt qu'une nuisance) est caractéristique de son art : puisqu'il n'existe pas de foncteur-fibre privilégié, il faut les prendre tous. Autrement dit, il faut prendre au sérieux la **catégorie** des foncteurs-fibres : les divers foncteurs-fibres sur une catégorie galoisienne fixée C s'organisent en un **groupoïde** profini (en un sens évident) non vide connexe : tout morphisme entre foncteurs-fibres est inversible et les foncteurs-fibres sont tous isomorphes l'un à l'autre, comme déjà indiqué. Ceci montre que le choix d'un foncteur-fibre ω détermine une équivalence du groupoïde des foncteurs-fibres sur celui des G-torseurs à droite, G étant comme ci-dessus le groupe des automorphismes de ω et G-torseur à droite désignant tout espace profini non vide, muni d'une action à droite continue, simple et transitive de G. Inversement, à tout groupoïde profini non vide connexe Φ, on peut associer la catégorie des systèmes locaux en ensembles finis sur Φ, i.e. des foncteurs $L : \Phi \longrightarrow Fin$, continus en ce sens que, pour tout objet ω de Φ, l'action fonctorielle du groupe profini $Aut_\Phi(\omega)$ sur l'ensemble fini $L(\omega)$ est continue (pour la topologie discrète de cet ensemble comme d'habitude). Cette catégorie est galoisienne et le théorème fondamental indiqué ci-dessus peut alors se reformuler sous une forme intrinsèque plus plaisante : les deux constructions que nous venons de mentionner sont quasi-inverses l'une de l'autre et forment une anti-équivalence de la 2-catégorie des catégories galoisiennes, où les 1-morphismes sont les foncteurs exacts, sur celle des groupoïdes profinis non vides connexes,

où les 1-morphismes sont les foncteurs continus. Sous cette forme, on voit que le cœur de la théorie est un **théorème de dualité** (un autre terme-clé dans la panoplie grothendieckienne), le théorème de **dualité galoisienne**.

La théorie ainsi obtenue a toute la souplesse désirable. C'est ainsi qu'en pratique, il peut y avoir avantage, dans l'étude d'une catégorie galoisienne "concrète" C, à n'utiliser que certains foncteurs-fibres. En fait, tout sous-groupoïde plein non vide du groupoïde des foncteurs-fibres de C convient, i.e. suffit évidemment à reconstituer la catégorie d'origine C, le cas où on n'utilise qu'un unique foncteur-fibre, seul cas considéré classiquement, étant un cas limite ; des considérations de symétrie par exemple, peuvent faire préférer le choix de tout un ensemble de "points-base", i.e. de foncteurs-fibres, plutôt qu'un seul.

Revenons à un corps k et à la catégorie Et_k des k-algèbres étales. On voit alors que la catégorie opposée est galoisienne et que toute extension algébriquement close de k en fournit un foncteur-fibre. Le théorème de dualité galoisienne ci-dessus redonne donc l'anti-équivalence de 2.3, le résultat central de la théorie de Galois version Grothendieck. La théorie des catégories galoisiennes que je viens d'esquisser contient donc cette dernière. Elle contient également la généralisation de cette théorie que constitue la théorie du groupe fondamental étale des schémas ; c'est en fait pour celle-ci que la théorie générale a été conçue à l'origine.

3.2 Groupe fondamental étale des schémas

Considérons un schéma S non vide connexe, supposé de plus localement nœthérien pour des raisons techniques, et notons Et_S la catégorie de ses revêtements étales, autrement dit la sous-catégorie pleine de la catégorie des S-schémas formée des schémas finis et étales au-dessus de S. Rappelons qu'un morphisme de schémas est dit étale s'il est localement de présentation finie, plat et non ramifié. On démontre que cette catégorie est galoisienne et que tout point géométrique $s : spec\,\Omega \longrightarrow S$ de S (i.e. tout point de S à valeurs dans un corps algébriquement clos Ω) en définit un foncteur-fibre $\omega_s : Et_S \longrightarrow Fin$ (associant au revêtement X de S l'ensemble des relèvements de s à X, ou encore (l'ensemble sous-jacent à) sa fibre en s, i.e. l'image inverse $s^{-1}X = spec\,\Omega \times_S X$). La théorie précédente fournit donc un groupe profini $\pi_1(S, s)$, le groupe fondamental (étale) de S en s, et une équivalence de catégories de Et_S sur la catégorie des $\pi_1(S, s)$-ensembles finis. Elle fournit même, plus intrinsèquement, un groupoïde profini non vide connexe $\Pi_1(S)$, le groupoïde fondamental étale de S, dont les objets sont les points géométriques de S et les flèches, les isomorphismes des foncteur-fibres correspondants, ainsi qu'une équivalence de Et_S sur la catégorie des $\Pi_1(S)$-systèmes locaux finis, comme ci-dessus. De plus, ces constructions sont fonctorielles en (S, s) et en S respectivement. Enfin, dans le cas où S est le spectre d'un corps k, on retrouve exactement la théorie de Galois-Grothendieck : les revêtements étales de $spec\,k$ sont les spectres des k-algèbres étales, le foncteur-fibre ω ci-dessus s'identifie au foncteur-fibre défini en 2.3 par l'extension algébriquement close Ω de k et le groupe fondamental $\pi_1(spec\,k, s)$ au groupe de Galois défini au même endroit.

Ajoutons que cette théorie du groupe fondamental joue un rôle extrêmement important en géométrie algébrique. Elle constitue la première étape pour une théorie de l'homotopie des schémas et est intimement liée à la théorie de la cohomologie étale.

3.3 Le cas d'un schéma normal

Nous venons de voir que la théorie du groupe fondamental étale, outre qu'elle est modelée sur la théorie de Galois classique, est directement liée à celle-ci dans le cas où le schéma S considéré est le spectre d'un corps. Il est une autre situation, importante et beaucoup plus générale, où l'interprétation en termes galoisiens ordinaires du groupe fondamental est immédiate, celle où le schéma S est normal. Cette situation se rencontre en particulier en arithmétique, où S peut être le spectre de l'anneau des entiers d'un corps de nombres ou une courbe lisse sur un corps fini.

Soient donc S un tel schéma normal (et localement nœthérien, non vide, connexe), donc intègre, k son corps des fractions et Ω une extension algébriquement close de k. Celle-ci définit un point géométrique s de $spec\, k$ et de S, d'où des groupes fondamentaux $\pi_1(spec\, k, s)$ et $\pi_1(S, s)$, dont le premier s'identifie, comme nous l'avons vu, au groupe de Galois $Gal(k)$ de k défini par Ω. Le morphisme de schémas $spec\, k \longrightarrow S$ fournit un homomorphisme de groupes profinis, de $Gal(k)$ dans $\pi_1(S, s)$. Le résultat est alors que cet homomorphisme est surjectif, exhibant ainsi le groupe fondamental $\pi_1(S, s)$ comme un quotient du groupe de Galois $Gal(k)$; il ne reste plus à la théorie de Galois ordinaire qu'à décrire ce quotient comme le groupe de Galois de l'extension K/k, où K est la réunion des extensions finies séparables de k dans Ω qui sont non ramifiées en tout point de S.

Comme le fait remarquer Grothendieck, cet énoncé peut servir de définition *ad hoc* du groupe fondamental dans le cas considéré ; c'est bien ainsi qu'avaient procédé les arithméticiens et géomètres, Lang et Serre notamment, avant l'avènement de la théorie générale.

3.4 La suite exacte fondamentale d'homotopie

Considérons un corps k et un k-schéma non vide S, de type fini et géométriquement connexe. Choisissons une clôture algébrique Ω de k et un k-point $s : spec\, \Omega \longrightarrow S$ de S à valeurs dans Ω. Notons $\bar{S}$ le schéma $S \otimes_k \Omega$, $\bar{s}$ le point de $\bar{S}$ à valeurs dans Ω déduit de s et $Gal(k)$ le groupe de Galois de k défini par Ω.

On obtient alors la suite exacte d'homotopie de Grothendieck

$$* \longrightarrow \pi_1(\bar{S}, \bar{s}) \longrightarrow \pi_1(S, s) \longrightarrow Gal(k) \longrightarrow *$$

issue du carré cartésien

$$\begin{array}{ccc} \bar{S} & \longrightarrow & S \\ \downarrow & & \downarrow \\ spec\, \Omega & \longrightarrow & spec\, k \end{array}$$

Cette suite exacte est véritablement fascinante : elle relie organiquement la géométrie à l'algèbre, voire à l'arithmétique (lorsque k est un corps global). Elle induit en effet une opération extérieure du groupe de Galois $Gal(k)$ (l'algèbre) sur le groupe fondamental géométrique $\pi_1(\bar{S}, \bar{s})$ (la géométrie), une opération extrêmement précieuse. Cette suite exacte et ce lien sont à la source des applications les plus profondes de la théorie à la géométrie arithmétique. Ils jouent un grand rôle également en géométrie anabélienne, comme

nous le verrons par la suite. Par exemple, tout choix d'un point rationnel de S permet de scinder la suite exacte et donc de relever l'opération extérieure de $Gal(k)$ sur le groupe fondamental géométrique en une vraie opération. La non-trivialité de la suite exacte est donc une obstruction à l'existence de points rationnels sur S, d'où un intérêt diophantien potentiel évident.

Remarquons qu'on a mieux en fait : la théorie fournit par fonctorialité une action du groupe profini $Gal(k)$ sur le groupoïde fondamental géométrique $\Pi_1(\bar{S})$, d'où les invariants précédents se déduisent d'après des propriétés générales des actions de groupe sur un groupoïde.

3.5 Le théorème d'existence de Riemann

Pour pleinement justifier le qualificatif "géométrique" accolé au groupe fondamental $\pi_1(\bar{S}, \bar{s})$ ci-dessus, il faut étudier ce qui se passe dans le cas où le corps de base k est le corps des complexes. Soit donc maintenant S un schéma localement de type fini sur le corps des complexes et S^{an} l'espace analytique sous-jacent. Grothendieck prouve alors un résultat remarquable et fondamental, généralisant les profondes découvertes de Riemann, qu'il nomme pour cette raison le théorème d'existence de Riemann : le foncteur image inverse par le morphisme naturel d'espace annelé de S^{an} vers S, de la catégorie des revêtements étales de S vers celle des revêtements analytiques finis non ramifiés de S^{an}, est une équivalence de catégories ; d'où il découle, si S est connexe, que, s étant un point de S^{an} et notant encore s le point de S à valeurs dans $\mathbb{C}$ correspondant, le groupe fondamental étale $\pi_1(S, s)$ s'identifie canoniquement au complété profini du groupe fondamental "ordinaire" $\pi_1(S^{an}, s)$.

Ce théorème apporte la justification demandée et parachève la relation géométrie-arithmétique évoquée ci-dessus. De plus, il permet de mettre en œuvre des techniques "transcendantes" dans l'étude du groupe fondamental étale et de calculer certains d'entre eux. Signalons que, même actuellement, on ne connait toujours pas de démonstration purement algébrique de certains des renseignements ainsi obtenus par voie analytique.

4 Catégories tannakiennes et motifs

4.1 Introduction

Nous abordons maintenant ce qui constitue sans doute le cœur de l'œuvre de Grothendieck, la théorie des motifs, qui, sans doute également, figure parmi ses travaux les plus profonds et les plus personnels. Née dans la clandestinité, la théorie est longtemps restée confidentielle, sans doute vu son caractère largement conjectural et même spéculatif, ne faisant l'objet que de rares publications, et bien maigres, vu l'ampleur du sujet ; Grothendieck en particulier n'a lui-même jamais eu, ou pris, l'occasion de publier quoi que ce soit sur ses méditations motiviques, à l'exception de quelques sybillines et éparses allusions. Cet état de choses n'a pas peu contribué à la réputation d'ésotérisme de la théorie. La situation a maintenant bien changé : après avoir été un privilège réservé à un petit nombre d'initiés, la théorie des motifs a peu à peu conquis de vastes espaces mathématiques, surtout à partir de la fin des années soixante-dix, et joue actuellement un rôle considérable dans toutes

les branches de la géométrie algébrique/arithmétique, et même au-delà. Les toute récentes médailles Fields attribuées à L.Lafforgue et à V.Voevodsky en sont une illustration.

Tout ce cercle d'idées se rattache de bien des manières au thème galoisien : la théorie des motifs peut par exemple être considérée, entre autres, comme une extension aux variétés de dimension arbitraire de la théorie de Galois, celle-ci, dans cette optique, concernant les variétés de dimension zéro. De plus, de même que Grothendieck a conçu sa théorie des catégories galoisiennes pour enchâsser sa version de la théorie de Galois et sa théorie du groupe fondamental, il va façonner, pour sa théorie des motifs, l'écrin approprié, la théorie des catégories tannakiennes, que l'on peut voir comme une variante linéaire et géométrique de celle des catégories galoisiennes. Précisons tout de suite que, partiellement inspirée de la classique théorie de dualité de Tannaka-Krein-Chevalley, qui lui vaut son nom, cette théorie des catégories tannakiennes, présentée dans la thèse de Saavedra Rivano, élève de Grothendieck, sera corrigée et complétée sur un point essentiel par Deligne.

Les idées galoisiennes précédemment esquissées de Grothendieck vont donc ressurgir et resservir lorsque celui-ci se lance dans ses gigantesques entreprises motiviques, à partir de 1964 au plus tard, comme l'indique une lettre fameuse envoyée à Serre par Grothendieck. Précisons qu'il conçoit d'emblée l'existence de catégories de motifs au-dessus de schémas de base quelconques, voire de bases plus générales encore, obéissant à un formalisme dérivé faisceautique des six opérations, objets que de nos jours on appelle souvent catégories de motifs mixtes, voire de faisceaux motiviques mixtes. Ici et pour simplifier, nous nous limiterons cependant à la situation la plus simple, mais déjà extraordinairement riche, celle des catégories de motifs (à coefficients rationnels) semi-simples (dits parfois purs) au dessus d'un corps de base F, motifs que nous nommerons motifs sur F, tout court, sans perdre de vue toutefois, qu'en isolant ainsi cette situation particulière de son contexte global, on perd inévitablement des éléments de structure et des informations. Ces motifs sont associés aux variétés projectives lisses sur F, comme nous le verrons ci-dessous. Pour ces catégories de motifs semi-simples sur un corps, Grothendieck a, entre autres choses, proposé une construction précise, en termes de cycles algébriques, construction cependant qui était alors, et reste, conjecturale, puisque sa validité dépend des fameuses conjectures standard, introduites et ainsi nommées par Grothendieck.

4.2 Les origines de la théorie

C'est la constatation de l'existence, pour les variétés algébriques (projectives lisses), de multiples théories cohomologiques — une notion que nous ne précisons pas ici, destinée à refléter et simplifier les propriétés arithmético-géométriques des variétés s'exprimant dans leurs cycles algébriques — et de leurs similitudes, qui va amener Grothendieck à l'idée des motifs. Comme dit Serre, on dispose de trop, en un sens, de théories cohomologiques, et de trop peu de relations entre celles-ci. Les plus connues de ces théories sont :

(1) pour F plongeable dans $\mathbb{C}$ et $j : F \longrightarrow \mathbb{C}$ un plongement, la **cohomologie de Betti**,

$$H_B(X) := H_{sing}(X^{an}_{\mathbb{C}}, \mathbb{Q})$$

(où $X_{\mathbb{C}}$ désigne la variété algébrique complexe définie par X et le plongement j, $X^{an}_{\mathbb{C}}$ la variété analytique associée et H_{sing} la cohomologie singulière),
une $\mathbb{Q}$-algèbre graduée de dimension finie munie d'une structure de Hodge,

(2) pour F quelconque, $\bar{F}$ une clôture algébrique de F et l un nombre premier distinct de la caractéristique de F, la **cohomologie l-adique**,

$$H_l(X) := \lim_{\overleftarrow{\nu}} H_{et}(\bar{X}, \mathbb{Z}/l^\nu \ mathbbZ) \otimes_{\mathbb{Z}_l} \mathbb{Q}_l$$

(où $\bar{X}$ désigne l'extension de X à $\bar{F}$ et H_{et} la cohomologie étale),
une $\mathbb{Q}_l$-algèbre graduée de dimension finie munie d'une action continue du groupe de Galois $Gal(\bar{F}/F)$,

(3) pour F de caractéristique nulle, la **cohomologie de de Rham**,

$$H_{DR}(X) := \mathbb{H}\,(X, \Omega^{\cdot}_{X/F})$$

(où $\Omega^{\cdot}_{X/F}$ désigne le complexe de de Rham de X sur F et $\mathbb{H}$ l'hypercohomologie),
une F-algèbre graduée filtrée de dimension finie,
où, dans chaque cas, X désigne une F-variété projective lisse arbitraire.

Un point important est que ces divers invariants associés à X ne sont pas indépendants ; ils sont liés par des isomorphismes, dits de comparaison, fonctoriels en X. On dispose par exemple, lorsque $F = \mathbb{Q}$ disons, des isomorphismes de de Rham (de Serre en fait)

$$H_{DR}(X) \otimes \mathbb{C} \longrightarrow H_B(X) \otimes \mathbb{C}$$

et d'Artin-Grothendieck

$$H_l(X) \longrightarrow H_B(X) \otimes \mathbb{Q}_l$$

qui montrent déjà que ces cohomologies ont même dimension.

Il reste cependant des similitudes inexpliquées et des problèmes ouverts. Par exemple, toute correspondance algébrique sur X induit des endomorphismes sur la cohomologie, donc des polynômes caractéristiques ; ceux-ci devraient être à coefficients rationnels et indépendants de la théorie cohomologique utilisée, ce que l'on ne sait pas prouver en général.

4.3 L'explication motivique et le lien avec Galois

La remarque de Grothendieck est que, dans cette situation, une explication naturelle serait fournie par l'existence d'une **théorie cohomologique universelle**, factorisant toutes les autres via des foncteurs dits de réalisation T_B, T_l, T_{DR},... , la catégorie dans laquelle cette théorie prend ses valeurs étant "par définition" la catégorie des motifs (sur F), que nous notons $\mathcal{M}_F$. On a donc un foncteur canonique de cohomologie motivique

$$h : \mathcal{V}_F^{op} \longrightarrow \mathcal{M}_F$$

où $\mathcal{V}_F$ désigne la catégorie des F-variétés projectives lisses.

Cette idée peut sembler naïve, exagéremment optimiste et même dépourvue de tout fondement. Mais Grothendieck a certainement de bonnes raisons d'y croire, qui n'appartiennent qu'à lui sans doute, en tout cas à l'époque. Quoi qu'il en soit, il poursuit son exploration avec opiniâtreté. L'étude des propriétés cohomologiques, déjà connues ou conjecturales, des variétés algébriques, lui suggère pour ses catégories motiviques $\mathcal{M}_F$, des structures et des

propriétés très riches : la catégorie $\mathcal{M}_F$ devrait être abélienne, $\mathbb{Q}$-linéaire et semi-simple ; elle devrait posséder un produit tensoriel naturel $\otimes$ (provenant du produit cartésien dans $\mathcal{V}_F$), des isomorphismes d'associativité et de commutativité, un objet unité $\mathbb{Q}(0)$ (la cohomologie motivique $h(*)$ de la F-variété $spec(F)$) et des Hom internes, le tout relatif à ce produit tensoriel ; elle devrait posséder de plus un objet inversible spécifique $\mathbb{Q}(1)$, baptisé motif de Tate, et provenant de la cohomologie motivique de la droite projective ; tous ses objets devraient être de longueur finie et tous ses Hom de dimension finie sur $\mathbb{Q}$; enfin $\mathcal{M}_F$ devrait être engendrée par ces opérations à partir de (l'image de) $\mathcal{V}_F$. Tout ceci donne une image assez précise de cette catégorie des motifs $\mathcal{M}_F$.

Ces structures et ces propriétés sont déjà bien familières. Outre évidemment la catégorie des $\mathbb{Q}$-espaces vectoriels de dimension finie, celle, plus généralement, des représentations rationnelles de dimension finie Rep_G d'un $\mathbb{Q}$-schéma en groupe G (que l'on peut supposer affine) est abélienne, $\mathbb{Q}$-linéaire, semi-simple si G est pro(réductif), munie d'un produit tensoriel, et les autres structures et propriétés de $\mathcal{M}_F$ (graduation, polarisation dont nous n'avons pas parlé,...) s'interprètent plus ou moins aisément en termes de (schémas en) groupes. La catégorie des motifs $\mathcal{M}_F$ serait-elle donc la catégorie des représentations d'un schéma en groupes approprié ? C'est là une idée plutôt fantastique, et qui crée un lien avec les théories de Galois et de Tannaka-Krein-Chevalley. Ce qui est plus fantastique encore est que la réponse est essentiellement positive : choix d'un "point-base" mis à part, il peut être nécessaire d'étendre le corps $\mathbb{Q}$ des coefficients pour "voir" le groupe.

La question se pose donc à Grothendieck de caractériser les catégories de la forme Rep_G, où G est un schéma en groupes affine sur un corps arbitraire k, et de reconstituer le groupe G à partir de sa catégorie de représentations rationnelles de dimension finie Rep_G, problèmes tout à fait analogues à ceux que traite la théorie des catégories galoisiennes. Les réponses constituent la théorie des catégories tannakiennes.

4.4 Catégories et dualité tannakiennes

Soient k un corps commutatif et $\mathcal{T}$ une catégorie abélienne k-linéaire munie d'un foncteur $\otimes : \mathcal{T} \times \mathcal{T} \longrightarrow \mathcal{T}$, dit produit tensoriel, d'isomorphismes d'associativité et de commutativité et d'un objet unité $\mathbf{1}$ pour ce produit. Grothendieck dit que $\mathcal{T}$ est **tannakienne** si :

(1) $\mathcal{T}$ admet des "Hom internes", i.e., pour tout couple d'objets (X,Y) de $\mathcal{T}$ est défini, par l'isomorphisme d'adjonction

$$Hom(Z, \mathcal{H}om(X,Y)) = Hom(Z \otimes X, Y)$$

fonctoriel en Z, un objet $\mathcal{H}om(X,Y)$ de $\mathcal{T}$, fonctoriel en (X,Y), et ce de sorte que, pour tous objets X_1, X_2, Y_1 et Y_2 de $\mathcal{T}$, le morphisme naturel

$$\mathcal{H}om(X_1,Y_1) \otimes \mathcal{H}om(X_2,Y_2) \longrightarrow \mathcal{H}om(X_1 \otimes X_2, Y_1 \otimes Y_2)$$

soit un isomorphisme,

(2) le morphisme naturel $k \longrightarrow End(\mathbf{1})$ est un isomorphisme,

(3) $\mathcal{T}$ possède un foncteur-fibre à valeurs dans un k-schéma S **non vide**, "foncteur-fibre" désignant ici tout foncteur $\omega : \mathcal{T} \longrightarrow \mathcal{Q}coh_S$ k-linéaire, exact et compatible aux produits tensoriels, de $\mathcal{T}$ vers la catégorie $\mathcal{Q}coh_S$ des $\mathcal{O}_S$-modules quasi-cohérents (ω étant alors

nécessairement fidèle et à valeurs dans la sous-catégorie pleine des $\mathcal{O}_S$-modules localement libres de rang fini).

Veuille le lecteur nous pardonner quelques imprécisions dans cette définition

Notons que la condition (2), indispensable mais qui faisait défaut dans la version originale exposée par Saavedra, a été mise en évidence par Deligne, qui est ainsi parvenu à compléter la théorie en démontrant le théorème fondamental énoncé ci-dessous. Notons également que, dans la condition (3), on peut supposer que S est le spectre d'une extension de k.

Si G est un schéma en groupes sur k, on vérifie que la catégorie Rep_G est tannakienne, le foncteur d'oubli de celle-ci vers la catégorie des k-espaces vectoriels étant un foncteur-fibre. La réciproque n'est vraie que localement. Plus précisément, le théorème de **dualité tannakienne** prend la forme suivante : les divers foncteurs-fibres sur une catégorie tannakienne $\mathcal{T}$ s'organisent en une **catégorie fibrée** $\Phi \longrightarrow Sch/k$ au-dessus de la catégorie Sch/k des k-schémas ; cette catégorie fibrée est même en fait une **gerbe affine** sur Sch/k munie de la topologie fpqc (il s'agit bien sûr d'une topologie au sens de Grothendieck). Autrement dit :

(1) Pour tout k-schéma S, les foncteurs-fibres sur $\mathcal{T}$ à valeurs dans S forment un groupoïde Φ_S,

(2) Pour tout morphisme $u : T \longrightarrow S$ de k-schémas et tout objet ω de Φ_S, on peut définir dans Φ_T, l'image inverse $u^*\omega$ de ω par u, fonctorielle en ω et compatible à la composition des morphismes de k-schémas,

(3) Pour tout k-schéma S et tout couple d'objets (ω_1, ω_2) de Φ_S, le foncteur $(u : T \longrightarrow S) \longmapsto Hom_{\Phi_T}(u^*\omega_1, u^*\omega_2)$ de la catégorie des S-schémas vers celle des ensembles, est un faisceau sur S pour la topologie fpqc,

(4) Ce faisceau est en fait représentable par un S-schéma affine,

(5) Les objets de Φ eux-mêmes, les foncteurs-fibres, sont de nature locale pour la topologie fpqc, i.e. ils vérifient la propriété de descente fpqc,

(6) Φ est localement non vide et localement connexe : il existe un k-schéma S non vide tel que Φ_S est non vide et, pour tout k-schéma S, deux objets quelconques de Φ_S sont localement isomorphes.

Inversement, à toute gerbe affine $\mathcal{G}$ sur k (i.e. sur Sch/k munie de la topologie fpqc), on peut associer la catégorie $Rep_{\mathcal{G}}$ de ses représentations, i.e. des (Sch/k)-foncteurs cartésiens de $\mathcal{G}$ vers le champ $\mathcal{Q}coh$ des modules quasi-cohérents sur les k-schémas, et cette catégorie est tannakienne.

Enfin ces deux constructions sont quasi-inverses l'une de l'autre et définissent une anti-équivalence de la 2-catégorie des catégories tannakiennes sur k vers celle des gerbes affines sur k : tel est le théorème fondamental de la théorie, un point-clé, démontré, comme nous l'avons déjà indiqué, par Deligne, étant la connexité locale mentionnée en (6).

La parenté de ce théorème de dualité tannakienne avec le théorème de dualité galoisienne de la précédente section saute aux yeux. On peut en particulier considérer la notion de gerbe comme une localisation de celle de groupoïde non vide connexe et la dualité tannakienne comme une variante linéarisée et localisée de la dualité galoisienne. La gerbe $\Phi \longrightarrow Sch/k$ associée comme ci-dessus à une catégorie tannakienne $\mathcal{T}$ mérite bien le nom de gerbe de Galois de $\mathcal{T}$.

Le résultat peut se formuler plus concrètement : si $\mathcal{T}$ est une catégorie tannakienne sur k, le choix d'un foncteur-fibre à valeurs dans un k-schéma S non vide définit un k-schéma en groupoïdes opérant sur S, affine et fidèlement plat sur S^2 — un objet moins intrinsèque

mais moins encombrant qu'une gerbe —, en termes duquel on peut reconstituer $\mathcal{T}$ (et donc la gerbe $\Phi \longrightarrow Sch/k$ tout entière), comme étant la catégorie des représentations de ce groupoïde. Plus concrètement encore, supposons que $\mathcal{T}$ admette un foncteur-fibre ω à valeurs dans k (i.e. $spec\, k$) lui-même, auquel cas $\mathcal{T}$ est dite neutre (et neutralisée par ω) ; le foncteur en groupes des k-automorphismes (compatibles aux produits tensoriels) de ω est alors représentable par un k-schéma en groupes affine G et $\mathcal{T}$ s'identifie par ω à la catégorie tannakienne Rep_G des représentations de G. Ce groupe mérite le nom de groupe de Galois de $\mathcal{T}$ ponctué par ω et, dans l'équivalence tannakienne ci-dessus, les catégories tannakiennes neutres sur k "correspondent" ainsi aux k-schémas en groupes affines.

La théorie que nous venons d'esquisser a trouvé bien des applications. Elle est en particulier récemment venue éclairer, à la suite de Deligne, la théorie de Galois différentielle. Mais voici le moment de revenir à la théorie pour laquelle elle fut initialement conçue.

4.5 Retour aux motifs

F étant un corps (commutatif) arbitraire, la catégorie $\mathcal{M}_F$ des motifs introduite précédemment, possède, conjecturalement, toutes les propriétés que l'on attend d'une catégorie tannakienne sur le corps $\mathbb{Q}$ des rationnels, y compris l'existence d'un foncteur-fibre : la cohomologie l-adique H_l en fournit un par exemple, pour tout nombre premier l distinct de la caractéristique de F. La théorie associe donc à F une $\mathbb{Q}$-gerbe affine, la gerbe de Galois motivique de F. Les objets de cette gerbe peuvent être vus comme les diverses théories cohomologiques définies sur $\mathcal{V}_F$. Cette catégorie tannakienne des motifs est neutre si, par exemple, F se plonge dans $\mathbb{C}$, car alors la cohomologie de Betti définie par un tel plongement la neutralise. C'est le cas en particulier si F est un corps de nombres et, plus particulièrement encore, si $F = \mathbb{Q}$, cas auquel nous allons nous limiter désormais, pour simplifier et fixer les idées ; d'ailleurs ce cas est tout spécialement intéressant pour l'arithmétique.

Grâce au foncteur-fibre H_B (disons), la catégorie $\mathcal{M}_{\mathbb{Q}}$ des motifs sur $\mathbb{Q}$ s'identifie donc à la catégorie des représentations rationnelles d'un $\mathbb{Q}$-schéma en groupes affine $Gal_{Mot,\mathbb{Q}}$, groupe des automorphismes tensoriels de H_B : c'est le groupe de Galois motivique du corps des rationnels. Ce $\mathbb{Q}$-schéma en groupes affine est réductif (ou pro-réductif), puisque $\mathcal{M}_{\mathbb{Q}}$ est semi-simple. Il devrait s'insérer dans une suite exacte

$$* \longrightarrow Gal_{Mot,\bar{\mathbb{Q}}} \longrightarrow Gal_{Mot,\mathbb{Q}} \longrightarrow Gal_{\mathbb{Q}} \longrightarrow *$$

à comparer avec la suite exacte d'homotopie de 3.4, et dans laquelle $\bar{\mathbb{Q}}$ désigne la clôture algébrique de $\mathbb{Q}$ dans $\mathbb{C}$, $Gal_{\mathbb{Q}}$ le groupe de Galois "ordinaire" associé, vu comme $\mathbb{Q}$-schéma en groupes constant, et enfin $Gal_{Mot,\bar{\mathbb{Q}}}$ le groupe de Galois motivique défini comme précédemment par la catégorie $\mathcal{M}_{\bar{\mathbb{Q}}}$ et la cohomologie de Betti. De plus, ce noyau $Gal_{Mot,\bar{\mathbb{Q}}}$ devrait être exactement la composante neutre de $Gal_{Mot,\mathbb{Q}}$. Noter que la projection sur $Gal_{\mathbb{Q}}$ provient de l'inclusion dans $\mathcal{M}_{\mathbb{Q}}$ de la sous-catégorie tannakienne engendrée par (la cohomologie motivique) des variétés de dimension zéro, sous-catégorie dite des motifs d'Artin. Noter également que, pour chaque nombre premier l, la cohomologie l-adique fournit une section sur $\mathbb{Q}_l$, $Gal_{\mathbb{Q}} \longrightarrow Gal_{Mot,\mathbb{Q}} \otimes \mathbb{Q}_l$, de cette suite exacte.

Ce groupe de Galois motivique $Gal_{Mot,\mathbb{Q}}$ est "énorme" mais peut être ramené à taille humaine : pour tout objet M de $\mathcal{M}_{\mathbb{Q}}$, la sous-catégorie tannakienne engendrée par M s'identifie par H_B à la catégorie des représentations rationnelles d'un $\mathbb{Q}$-schéma en groupes affine

Gal(M), quotient du gros groupe $Gal_{Mot,\mathbb{Q}}$ et nommé groupe de Galois (motivique) du motif M. Inversement, le groupe global $Gal_{Mot,\mathbb{Q}}$ peut être reconstruit comme la limite projective filtrante des Gal(M) pour M variable dans $\mathcal{M}_{\mathbb{Q}}$. Ce groupe de Galois Gal(M), étant de type fini, est en fait, dans un langage plus terre à terre, un groupe algébrique linéaire défini sur $\mathbb{Q}$, réductif de surcroît. C'est aussi l'image de la représentation $Gal_{Mot,\mathbb{Q}} \longrightarrow GL(H_B(\mathrm{M}))$ associée à M par dualité tannakienne. Cette merveilleuse construction associe donc à tout motif (sur $\mathbb{Q}$) un objet des plus classiques et des plus sympathiques, un groupe réductif (sur $\mathbb{Q}$), son groupe de Galois motivique, un invariant clairement du plus haut intérêt, par lequel s'amorcent les relations avec l'analyse harmonique et la théorie des formes automorphes.

4.6 Cycles algébriques et retour à Galois

Les idées exposées dans cette section — et les précédentes — présentent un caractère incontestablement galoisien, nous croyons l'avoir montré amplement. On pourrait se croire néanmoins, au terme de ce long périple avec Grothendieck, installé sur des cimes tellement élevées que le lien avec la théorie de Galois originelle se soit perdu en cours de route. Il n'en est rien cependant, et redescendre dans la vallée est très facile.

Considérons en effet une variété projective lisse X, définie sur $\mathbb{Q}$ toujours pour fixer les idées. Un ingrédient essentiel de toute théorie cohomologique au sens où nous l'utilisons ici, est la possibilité d'associer une classe de cohomologie à tout cycle algébrique sur X. En termes motiviques, on trouve, pour un tel cycle supposé défini sur $\mathbb{Q}$, un morphisme de motifs, du motif unité $\mathbb{Q}(0)$ vers le motif $\mathrm{M} = h^{2d}(X)(d)$, où d désigne la codimension du cycle considéré et (d) signifie qu'on a tensorisé d fois le motif $h^{2d}(X)$ par le motif de Tate $\mathbb{Q}(1)$. Autrement dit (dualité tannakienne), on trouve une classe dans $H_B^{2d}(X)(d)$ **invariante** par le groupe de Galois du motif M. Autrement dit encore, l'existence de ces cycles sur X "contraint" ce groupe de Galois, de même qu'en théorie de Galois ordinaire, l'existence de relations algébriques entre les racines d'un polynôme séparable contraint son groupe de Galois. De plus, une généralisation de la construction précédente aux associés tensoriels de M devrait conjecturalement fournir suffisamment de classes invariantes pour que le groupe de Galois de M soit précisément le plus petit sous-groupe de $GL(H_B^{2d}(X)(d))$ laissant invariantes lesdites classes. Sous ce rapport, Grothendieck apparait donc comme le continuateur direct de Galois : ce que celui-ci avait fait pour les variétés de dimension zéro, Grothendieck le réalise pour les variétés de dimension arbitraire ; il poursuit l'analyse des équations algébriques que Galois avait initiée, en passant de une à plusieurs inconnues.

4.7 Pour conclure

Nous devons arrêter ici cette rapide évocation des motifs ; encore n'avons nous pu mentionner nombre de questions, ne serait-ce que le yoga des poids, les liens de la théorie avec les jacobiennes et variétés abéliennes, avec les fonctions L, avec les périodes, avec les formes automorphes et le programme de Langlands, etc... Il était de toutes façons impossible de rendre ici justice à l'extraordinaire richesse de la philosophie motivique.

On pourra objecter que tout ceci n'est qu'une suite de rêveries fumeuses. Mais s'il est vrai que la théorie ne fonctionne que grâce à une série de "miracles", la grande force de la création motivique de Grothendieck réside dans sa profonde cohérence, dans ses pouvoirs

explicatif et prédictif heuristiques, sans même parler de sa suprême beauté. D'ailleurs, **Grothendieck revendique la liberté de rêver**, y compris en mathématiques, et même de prendre ses rêves pour des réalités, en quelque sorte, ce qui ne lui a pas mal réussi. Quoi qu'il en soit, après n'avoir été que fort peu cultivée pendant une quinzaine d'années, la théorie est sortie de l'ombre désormais et a connu d'innombrables développements. Domaine de recherche en pleine activité, elle a été étayée de multiples manières par les travaux de nombre de mathématiciens et a bénéficié des éclairages nouveaux apportés par la théorie des faisceaux pervers et des t-structures, celle des D-modules, par le programme de Langlands, par la K-théorie, etc... Aujourd'hui comme hier, elle joue le rôle d'un guide heuristique, d'une sûreté et d'une fécondité sans failles. Les doutes ne sont plus guère de mise de nos jours : le rêve, la vision motiviques de Grothendieck, sans doute l'une des plus ambitieuses et audacieuses entreprises de l'histoire de notre discipline, sont devenus une réalité mathématique incontestable, quoique toujours enveloppée de mystère.

5 Géométrie anabélienne, théorie de Galois-Teichmüller et dessins d'enfant

5.1 Introduction

Nous donnons, pour terminer cet article, quelques brèves, et malheureusement nécessairement vagues, indications sur quelques-unes des idées que Grothendieck a commencé à développer dans les années soixante-dix. Il s'agit, là encore, d'un programme de recherche extrêmement ambitieux, articulé autour de plusieurs thèmes, parmi lesquels se trouvent ceux qui figurent dans le titre de la présente section.

Ce qui fait l'unité de ces thèmes, et justifie leur évocation dans ce texte, est leur caractère galoisien prononcé et totalement explicite. Ces recherches représentent un retour de Grothendieck sur son passé, un retour en fait à sa théorie du groupe fondamental étale, une théorie décidément d'une extraordinaire richesse et profondeur, dont les potentialités étaient restées insoupçonnées jusqu'alors. Tous les thèmes précités sont en effet issus de ladite théorie, et même, plus spécifiquement, de la suite exacte fondamentale de 3.4 et de l'action extérieure du groupe de Galois qui s'en déduit. C'est celle-ci notamment qui fascine tout particulièrement Grothendieck ; elle offre en effet de nouveaux moyens d'étude, et des variétés, et des groupes de Galois eux-mêmes. Ceux-ci viennent enrichir et renouveler l'arsenal de la géométrie algébrique. Les géomètres avaient jusqu'alors abondamment étudié et exploité des représentations galoisiennes **linéaires** ; par constrate, ce qui fait la richesse des nouvelles représentations proposées par Grothendieck, est leur caractère fortement **non abélien**.

Donnons maintenant quelques détails sur les thèmes en question.

5.2 Géométrie algébrique anabélienne

Soit k un corps de type fini et de caractéristique nulle ; fixons une clôture algébrique de k et notons Γ_k le groupe de Galois correspondant. Nous aurons à considérer des k-schémas de type fini S, non vides et géométriquement connexes, que nous nommerons variétés

simplement, ainsi que leurs groupes fondamentaux étales, arithmétiques ou mixtes, $\pi_1(S)$, et géométriques, $\pi_1(\bar{S})$, dans lesquels nous pourrons, comme nous venons de le faire, négliger la mention des points-base.

Le point de départ de la géométrie anabélienne est la conviction acquise par Grothendieck, que "beaucoup" de variétés sont déterminées par la suite exacte de 3.4 correspondante, c'est-à-dire par leur groupe fondamental arithmétique vu comme groupe profini au-dessus du groupe de Galois Γ_k. Cette conviction l'amène à conjecturer l'existence d'une classe $\mathcal{A}nab_k$ de k-variétés, dites anabéliennes, ayant les propriétés suivantes : pour X et Y dans $\mathcal{A}nab_k$,

(1) l'application naturelle de l'ensemble des k-isomorphismes de X dans Y, dans celui des isomorphismes de groupes profinis augmentés vers Γ_k, de $\pi_1(X)$ dans $\pi_1(Y)$, à conjugaison près par $\pi_1(\bar{Y})$, est bijective,

(2) l'application naturelle de l'ensemble des k-morphismes dominants de X dans Y, dans celui des morphismes de groupes profinis augmentés vers Γ_k, de $\pi_1(X)$ dans $\pi_1(Y)$, dont l'image est ouverte et à conjugaison près par $\pi_1(\bar{Y})$, est bijective,

(3) si X est propre et de dimension 1, l'application naturelle de l'ensemble des k-points de X dans celui des sections (continues) de l'augmentation de $\pi_1(X)$ vers Γ_k, à conjugaison près par $\pi_1(\bar{X})$, est bijective.

Grothendieck formule également des variantes absolues et birationnelles de cette conjecture anabélienne fondamentale. Plus spécifiquement, il est convaincu que, parmi les courbes lisses géométriquement connexes, celles qui sont anabéliennes sont les courbes hyperboliques, i.e. celles dont la caractéristique d'Euler-Poincaré est strictement négative.

Ces conjectures ont suscité un grand engouement et des travaux enthousiastes. Notons en particulier que le cas des courbes hyperboliques a été réglé par l'école japonaise, A.Tamagawa et Mochizuki en tête, et que la conjecture anabélienne birationnelle en dimension zéro avait été anticipée dans le cas des corps globaux par Neukirch et Uchida, le même résultat ayant été ensuite obtenu pour tous les corps de type fini et de caractéristique nulle par Pop.

La situation est donc bien clarifiée en dimensions 0 et 1. Signalons qu'en revanche, la définition à adopter pour les variétés anabéliennes en dimension plus grande est loin d'être évidente.

5.3 Théorie de Galois-Teichmüller et dessins d'enfant

A côté des courbes hyperboliques, les espaces modulaires $\mathcal{M}_{g,n}$, paramétrant, pour tous entiers naturels g et n, les courbes n-pointées projectives lisses connexes de genre g, ici sur le corps $\mathbb{Q}$ des rationnels, devraient être des exemples particulièrement remarquables de variétés anabéliennes. Liées entre elles par certaines opérations géométriques naturelles, elles forment, elles ou plutôt leurs groupoïdes fondamentaux géométriques, ce que Grothendieck nomme la **tour de Teichmüller**. Une propriété étonnante de cette tour, conjecturée par Grothendieck, est qu'elle devrait pouvoir se reconstituer à partir de ses deux premiers étages seulement, formés des $\mathcal{M}_{g,n}$ de dimension $d = 3g - 3 + n$ inférieure à 2.

En fait de variétés, ces $\mathcal{M}_{g,n}$ sont en réalité des objets d'une nature sensiblement plus générale et subtile, des champs algébriques, mais peu nous importe ici. La théorie du groupe fondamental étale s'étend en effet au contexte de ces champs algébriques ; on obtient donc en particulier, une action extérieure du groupe de Galois absolu $\Gamma = \Gamma_{\mathbb{Q}}$ sur les groupes fondamentaux géométriques des $\mathcal{M}_{g,n}$, qui ne sont autres que les compactifiés profinis $\hat{\mathrm{T}}_{g,n}$

des classiques groupes de Teichmüller $\mathrm{T}_{g,n}$. Ce qui fascine ici Grothendieck, est que cette action, compatible aux opérations naturelles reliant les $\mathcal{M}_{g,n}$ entre eux, est fidèle. Le problème est alors de tenter de caractériser l'image de Γ dans le groupe des automorphismes de la tour, et même dans celui de la sous-tour $(\hat{\mathrm{T}}_{0,n})$ de genre zéro, sur laquelle l'action de Γ est déjà fidèle, ce qui fournirait une description purement géométrique, voire combinatoire, du mystérieux groupe Γ.

Mais il y a plus remarquable encore : le fameux théorème de Belyi, largement anticipé par Grothendieck, montre qu'est déjà fidèle l'action extérieure de Γ sur le groupe $\hat{\mathrm{T}}_{0,4}$, un groupe profini libre sur deux générateurs, groupe fondamental du tout premier $\mathcal{M}_{0,4}$ de ces espaces modulaires, celui qui paramètre les droites projectives munies d'un quadruple de points distincts, isomorphe par le birapport à la droite projective sur $\mathbb{Q}$ privée des trois points 0, 1 et ∞. Là encore se pose le problème de caractériser l'image de Γ dans le groupe des automorphismes de $\hat{\mathrm{T}}_{0,4}$. De plus Grothendieck a proposé, toujours via le théorème de Belyi, une surprenante interprétation combinatoire des courbes algébriques projectives lisses sur $\bar{\mathbb{Q}}$, par des objets qu'il appelle cartes orientées finies, ou dessins d'enfant — il s'agit de certains 1-complexes finis convenablement plongés dans une surface topologique orientée compacte connexe — , interprétation qui corrobore l'espoir d'une description purement combinatoire, indépendante de la théorie des nombres algébriques, du groupe de Galois des rationnels.

5.4 L'apothéose de la théorie de Galois

Les conceptions de Grothendieck sur la géométrie anabélienne vont beaucoup plus loin que ce que nous avons evoqué ci-dessus. Ce qu'il envisage très rapidement, c'est, au-delà des variétés anabéliennes, la possibilité de reconstituer tout schéma de type fini sur $\mathbb{Q}$ à partir d'un diagramme convenable de groupes fondamentaux, puis de s'affranchir de la restriction à la caractéristique zéro et traiter les schémas de type fini sur $\mathbb{Z}$ lui-même, pour finalement parvenir, par des passages à la limite appropriés, aux schémas les plus généraux, voire aux champs et autres multiplicités algébriques...

En définitive, comme le souligne Grothendieck, la géométrie anabélienne devrait bien constituer une nouvelle approche de la géométrie algébrique, aboutissant à une description des schémas, plus précisément de leurs topos étales, en termes de diagrammes de groupes profinis. Il s'agit bien d'une nouvelle traduction d'un univers géométrique, celui des schémas, en un univers algébrique, celui des groupes profinis, prolongeant celle qu'avait effectuée Galois en son temps, d'un **aboutissement ultime** de la théorie de Galois, selon les propres termes de Grothendieck.

Un des traits les plus remarquables de ces nouvelles recherches de Grothendieck est que, si elles constituent incontestablement un renouvellement spectaculaire de sa pensée, elles ne sont en aucun cas une rupture avec son passé mathématique, dans lequel, au contraire, elles s'enracinent profondément ; l'univers anabélien et l'univers motivique devraient s'éclairer et se compléter mutuellement. Grothendieck entrevoit en effet dans la géométrie anabélienne, de nouvelles opportunités pour appréhender *la réalité encore mystérieuse des motifs* : des liens subtils devraient exister entre la nouvelle théorie, essentiellement non abélienne, et la théorie des motifs, distinctement abélienne.

Il semble bien que, là encore, Grothendieck ait vu juste : n'est-il pas permis de voir, dans l'identité, non encore prouvée mais fortement soupçonnée, entre le groupe de Grothendieck-

Teichmüller introduit par Drinfeld — il s'agit ici de la version $\mathbb{Q}$-algébrique pro-unipotente —, et, en substance, le groupe de Galois de la catégorie des $\mathbb{Q}$-motifs mixtes de Tate non ramifiés au-dessus de $spec\,\mathbb{Z}$, une première confirmation de la vision nouvelle ?

" Car il ne suffit pas que le monde soit beau
Encore faut-il daigner s'en réjouir... "

Alexandre Grothendieck

Les programmes de classification des variétés algébriques complexes : 1970-2000

Frédéric Campana[1]
Université Nancy 1
Département de Mathématiques
BP 239
F-54506.Vandoeuvre-les-Nancy.Cédex

6 Introduction

De nombreux domaines de la géométrie ont connu durant les deux dernières décennies des avancées importantes : la topologie différentielle en dimension quatre (invariants de Donaldson et de Seiberg-Witten ; la géométrie symplectique (courbes pseudo-holomorphes de Gromov et invariants de Gromov-Witten). La géométrie arithmétique également, avec la solution des deux conjectures fondamentales de Mordell et de Taniyama-Weil.

La géométrie algébrique entretient des relations étroites avec ces domaines de la géométrie, et leur sert souvent de motivation ou de source d'exemples. La géométrie symplectique en dimension (réelle) quatre tente d'obtenir une classification analogue à celle des surfaces projectives complexes . La théorie des courbes pseudo-holomorphes de M.Gromov est une version symplectique, dans le cas des courbes, des variétés de Chow en géométrie algébrique. Le cas particulier des courbes rationnelles pseudo-holomorphes étend la théorie de Mori, brièvement décrite ci-dessous. Enfin, la conjecture de Mordell n'est autre qu'une équivalence entre propriétés arithmétiques et géométriques (dans le sens qu'on lui donne ci-dessous : la positivité du fibré canonique) pour les courbes algébriques complexes définies sur un corps de nombres. Cette conjecture a d'ailleurs une version géométrique concernant les surfaces algébriques complexes, dans laquelle $\mathbb{Q}$ est remplacé par le "corps de fonctions" d'une courbe complexe. Finalement, les courbes modulaires de la conjecture de Taniyama-Weil sont les espaces de modules de courbes elliptiques munies de structures de niveau.

[1] campana@iecn.u-nancy.fr

Ces deux dernières décennies ont aussi vu apparaitre, par l'intermédiaire de la symétrie miroir et de la cohomologie quantique des relations entièrement nouvelles et trés profondes entre géométries différentielle et algébrique (sans oublier bien sûr la physique des champs, qui en a fourni certaines motivations).

On se propose ici de décrire succinctement certains des progrès réalisés depuis 1970 dans la classification des variétés projectives complexes, progrès qui ont bouleversé les perspectives et les méthodes de cette discipline.

Cette description tente d'éviter, autant que possible, le recours à un langage technique. Quelques préliminaires sont malheureusement nécéssaires pour transmettre au non-spécialiste (auquel ce texte est exclusivement destiné) une intuition géométrique minimale des variétés algébriques. On se basera pour cela sur l'analogie avec certaines notions de base de la géométrie différentielle, plus intuitives que celles de l'algèbre. Il nous a malheureusement semblé impossible, pour de nombreuses raisons, de faire un exposé suivant le développement historique. Par manque de compétence, certains thèmes importants ne sont pas traités : cohomologie quantique et symétrie miroir entre autres.

7 Les variétés projectives complexes

Les objets d'étude sont les variétés complexes connexes compactes X qui sont **projectives**, c'est-à-dire : admettent un plongement dans un espace projectif complexe $\mathbb{P}_N(\mathbb{C})$ (l'existence d'un tel plongement est, contrairement à l'analogue différentiable, une restriction très forte sur X, à la fois topologique et analytique).

Un théorème dû à Chow affirme qu'une telle X est alors définie par des équations polynomiales homogènes. Ceci permet d'étudier X par des méthodes soit analytiques, soit algébriques. Nombre de résultats essentiels ne sont d'ailleurs à ce jour accessibles que par une seule de ces méthodes. Nous en verrons un exemple typique avec la méthode de Mori de construction de courbes rationnelles, basée sur la "réduction modulo p" et le morphisme de Frobenius.

Le développement historique a procédé par dimension croissante de X. (Il s'agit de la dimension complexe, moitié de la dimension réelle, et notée n). On dit que X est une **courbe** (resp. une **surface**) si $n = 1$ (resp. 2). Une telle "courbe" est une surface réelle, et une telle "surface" est donc de dimension réelle quatre.

Anticipant sur l'évolution historique, nous noterons que ces recherches ont fait apparaitre une trichotomie fondamentale, constituée de 3 types de géométries que nous qualifierons ici de "sphérique", (Ricci-)"plate" et "hyperbolique" (ces termes sont destinés à fournir une idée intuitive, basée sur le signe de (l'opposée de) la courbure de Ricci ; ils ne sont pas standards).

Ces géométries sont définies par des notions adéquates de signe du **fibré canonique** $K_X := \det(\Omega^1_X)$ de X, déterminant du fibré cotangent complexe de X.

Ces propriétés de positivité gouvernent la géométrie (et l'arithmétique) de X, ceci à un point tel que l'on peut sans simplification trop abusive décrire la géométrie algébrique (ou arithmétique) comme étant la traduction en termes géométriques (ou arithmétiques) des propriétés de positivité (ou de négativité) du fibré canonique. Nous illustrerons cette assertion à plusieurs reprises ci-dessous, distinguant trois périodes de classification :

1. Avant 1970 : Courbes et Surfaces, mise en évidence de la trichotomie.
2. 1970-1980 : Le programme d'Iitaka-Moishezon, la composante "plate" en dimension arbitraire.
3. 1980 - 2000 : La dimension 3 et le programme des modèles minimaux, la construction de la composante "sphérique" en dimension 3.

8 Les Courbes

Les **courbes** sont apparues dans divers contextes : géométrie plane, problèmes diophantiens et intégrales abéliennes. Leur structure topologique et analytique a été graduellement élucidée au cours du 19^e siècle, sous l'influence des travaux de Riemann. Ces travaux ont fait apparaitre la trichotomie fondamentale entre géométries sphérique, plate et hyperbolique.

La structure différentiable d'une courbe complexe X est décrite par un unique invariant topologique : son **genre** $g = g(X)$, qui est un entier positif ou nul. Les trois classes du tableau sont simplement caractérisées par le signe du fibré canonique (c'est-a-dire ici l'opposé du signe de la courbure d'une métrique sur X, si cette courbure est constante). La concordance du signe du fibré canonique et des valeurs appropriées du genre traduit une relation entre propriétés analytiques complexes et topologie (ou structure différentiable).

g	X	structure $\mathcal{C}^\infty$	degré de K_X	p_m	
0	$\mathbb{P}_1(\mathbb{C})$	S^2	-2	0	
1	$\mathbb{C}/\Lambda$ $\Lambda \simeq \mathbb{Z}^2$ réseau	$S^1 \times S^1$	0	1	
≥ 2	Δ/Γ Δ disque unité Γ cocompact sans point fixe	$\sharp_1^g(S^1 \times S^1)$	> 0	$(2m-1)(g-1)$ $(m \geq 2)$ g si $m = 1$	 $g = 2$

Figure 1. Les courbes

On remarque que la situation se complique lorsque l'on passe de la géométrie sphérique à la géométrie plate, puis hyperbolique : du point de vue topologique, mais aussi du point de vue de l'abondance des variétés (une seule, rigide, en genre zéro ; une unique famille à un paramètre en genre 1, la structure complexe étant décrite par l'invariant "modulaire" j) ; une infinité de familles paramétrées par des variétés de dimension croissante en genre 2 ou plus). Plus précisément :

Pour $g \geq 2$ fixé, la structure complexe de X dépend de $(3g-3)$ "modules" complexes, qui (après ajout des courbes stables) décrivent une variété projective complexe, l'espace de Deligne-Mumford, dont les variantes "marquées" jouent un rôle essentiel dans la construction des invariants de Gromov-Witten, par exemple.

L'objectif de la classification (qui semble n'être devenu apparent qu'assez récemment) consiste à classifier les variétés de type "pur" (soit "sphérique", soit "plat", soit "hyperbolique", la définition précise de ces notions étant partie du problème), et à décomposer les variétés arbitraires à l'aide de fibrations "canoniques" en leurs composantes de type "pur", ceci de telle sorte que cette décomposition varie holomorphiquement lorsque l'on déforme la structure complexe. Les familles de déformations sont repérées par des invariants topologiques ou différentiables discrets, les variétés d'une famille le sont (à isomorphisme, plus éventuelle structure additionnelle près) par un "module", point d'un ouvert d'une variété projective, dont le complémentaire est encore une telle sous-variété paramétrant des dégénérescences contrôlées.

9 Le signe du fibré canonique

Conformément à ce qui a été dit plus haut, cette trichotomie est, en fait, basée sur les propriétés de **positivité du fibré canonique**, positivité qui peut être définie de diverses façons, fournissant des notions voisines. On peut mentionner :

1. La notion **métrique**, (déjà considérée et étudiée par A. Lichnérowicz) est définie par (l'opposé de) la courbure de Ricci (si h est une métrique hermitienne sur le fibré canonique, sa courbure $\mathbf{\Theta}$ est par définition la $(1,1)$-forme $\mathbf{\Theta} := (1/2\pi i)\partial\overline{\partial}\log|h(v(x))|^2$, où v est une section (holomorphe) locale sans zéro du fibré canonique, c'est-à-dire une forme volume holomorphe (écrite $v(x) = f(x)dx_1 \wedge \cdots \wedge dx_n$ dans des coordonnées locales holomorphes $x = (x_1, \ldots, x_n)$, avec f holomorphe). La classe de cohomologie de la forme de courbure (qui est bien sûr d-fermée) ne dépend pas de la métrique choisie, et est entière (*i.e.* dans $H^2(X, \mathbb{Z})$) ; c'est l'opposée de la première classe de Chern $c_1(X)$ de X.
   On peut alors dire que notre variété X est à géométrie sphérique (resp. plate, resp. hyperbolique) si cette courbure est (en chaque point, pour une métrique adéquate sur le fibré canonique) négative (resp. nulle, resp. positive). (La géométrie différentielle considère le fibré dual, d'où le conflit de signes). En général, une variété n'a pas de métrique ayant une courbure de signe constant (le produit d'une courbe rationnelle par une courbe de genre 2 ou plus fournit un tel exemple).
   Par contre, il existe en toute dimension (non nulle) des variétés ayant un fibré canonique de signe donné. L'exemple le plus simple est celui des hypersurfaces lisses X de degré d de l'espace projectif de dimension $n+1$: si $d \leq n+1$ (resp. $d = n+2$, resp. $d \geq n+3$), alors K_X est négatif (resp. nul, resp. positif). On retrouve donc, du point de vue du rapport entre le signe du fibré canonique d'une part, l'abondance et la complexité d'autre part, une situation similaire à celle des courbes.
   Une variété "sphérique" est dite **de Fano**. (Le terme de "sphérique" se rapporte à la courbure, pas à la topologie : la droite projective est la seule variété projective complexe ayant la topologie d'une sphère (d'homologie), comme le montre immédiatement le second nombre de Betti). Ces variétés sont cependant un peu les analogues des sphères en géométrie différentielle.
2. La notion **numérique**, définie par les nombres d'intersection $K_X \cdot C$ du fibré canonique avec les courbes complexes compactes C tracées sur X.(Ces nombres d'intersection

peuvent être définis de nombreuses manières, l'une d'elles étant comme intégrale prise sur la courbe C de la forme de courbure. Cette intégrale est donc un nombre entier qui ne dépend que de la classe d'homologie du 2-cycle supporté par C). (Les courbes tracées sur X peuvent avoir des singularités : on intègre sur leur normalisation l'image réciproque de la forme de courbure).
On dira alors que K_X est numériquement négatif (resp. numériquement nul ; resp. numériquement positif) si tous les nombres $K_X \cdot C$ le sont, lorsque C est une courbe compacte arbitraire tracée sur X. (On obtient alors en principe les mêmes notions que ci-dessus : c'est vrai dans le cas numériquement nul (par la solution par S.T. Yau de la conjecture de Calabi), et conjecturé dans les autres cas. Cette conjecture peut d'ailleurs être réduite à la non-nullité du nombre d'auto-intersection du fibré canonique) : $(K_X)^n := \int_X \mathbf{\Theta}^{\wedge n} \neq 0$).

3. La notion **géométrique** basée sur la considération des **plurigenres $\mathbf{p_m}$** de X : le m-ième **plurigenre** de X, noté $p_m = p_m(X)$, est défini pour $m > 0$ entier, comme étant la dimension (finie, par compacité de X et théorème de Montel) de l'espace vectoriel complexe des sections (holomorphes) de $K_X^{\otimes m}$, où K_X est le **fibré canonique** de X.

Rappelons qu'une section v de $K_X^{\otimes m}$, m entier positif, n'est autre qu'une **m-forme volume holomorphe**, c'est-à-dire un tenseur holomorphe sur X exprimé localement sous la forme $v(x) = f(x)(dx_1 \wedge \cdots \wedge dx_n)^{\otimes m}$ dans des coordonnées locales holomorphes $x = (x_1, \ldots, x_n)$.

Si C est une courbe tracée sur X et v une telle section non nulle de $K_X^{\otimes m}$, avec $m > 0$, on peut aussi définir $K_X \cdot C$ comme étant N/m, où N est le nombre de points (comptés avec multiplicités) de C en lesquels v s'annule (pourvu que v ne s'annule pas identiquement sur C, condition satisfaite si C est en position assez générale). En particulier, si $K_X \cdot C < 0$ pour toute courbe C tracée sur X (cas de $X = \mathbb{P}^n(\mathbb{C})$, par exemple), alors $p_m = 0, \forall m > 0$.

Géométrie	SPHÉRIQUE	PLATE	HYPERBOLIQUE
Métrique (signe de Θ)	$\Theta < 0$ "variétés de Fano"	$\Theta = 0$	$\Theta > 0$
Numérique (pour toute courbe complexe C tracée sur X)	$K_X \cdot C < 0$	$K_X \cdot C = 0$	$K_X \cdot C > 0$
"Géométrique"	$\forall m > 0, p_m = 0$	$\forall m^{(*)} > 0, p_m = 1$	$\forall m^{(*)} > 0, p_m \sim Cm^n$ C constante > 0
Ex : hypersurfaces lisses de degré d dans $\mathbb{P}_{n+1}(\mathbb{C})$	$d \leq n+1$	$d = n+2$	$d \geq n+3$
$n = 1$	droites, coniques	courbes ellipitiques	$g = (d-1)(d-2)/2$
$n = 2$	plans, quadriques, cubiques	$K3$	$d \geq 5$
$n = 3$	$\mathbb{P}_3(\mathbb{C})$, quandriques cubiques, quatriques	quintiques	$d \geq 6$

(*) m suffisamment divisible

Figure 2 : La Trichotomie en dimension supérieure

Les notions métriques et numériques sont très précises et rigides, et ne peuvent être réalisées que sur de bons modèles birationnels (voir ci-dessous) de X, contrairement aux notions géométriques qui sont, tout comme les plurigenres, des invariants birationnels.

Néammoins, les analogues "géométriques" des classes "plates" et "hyperboliques" sont adéquates, bien que légèrement plus vastes que les classes d'équivalence birationnelle de leurs versions "métriques" ; par contre l'analogue "géométrique" de la notion "sphérique" est beaucoup trop faible pour fournir une bonne description de la structure globale de X (voir troisième période ci-dessous).

En général, les propriétés métriques fournissent des propriétés géométriques grâce à des théorèmes d'annulation, tel celui de Kodaira. On remonte des propriétés géométriques aux propriétés métriques en deux étapes beaucoup plus difficiles : d'abord par construction d'un "modèle minimal" de X, qui jouit de bonnes propriétés numériques, puis on passe du modèle minimal aux propriétés métriques par construction de métriques de Kähler-Einstein, grâce à des variantes du théorème d'Aubin-Calabi-Yau.

La difficile construction de tels "modèles minimaux" est malheureusement nécessaire dans de nombreux cas où l'on souhaite établir des propriétés birationnelles que l'on ne sait démontrer que sur ces "modèles minimaux".

10 Les Surfaces

L'étude des **surfaces** a été initiée par Picard et Poincaré dans la décennie 1890-1900, et poursuivie par l'école Italienne (représentée entre autres par Castelnuovo, Severi et Enriques). Cette étude culmine avec la classification publiée par Enriques et Severi en 1949. Leurs arguments ont été modernisés et rendus rigoureux par O. Zariski durant la période 1930-1950.

Une classification définitive et moderne, basée sur la (alors récente) théorie des faisceaux est publiée d'abord par Kodaira en 1960-63 (qui considère également les surfaces non projectives, rendues accessibles par le tout nouveau théorème de l'indice d'Atiyah-Singer), et indépendamment par I. Shafarevitch et ses collaborateurs.

Pour les surfaces, la classification comporte deux étapes : la première (inexistante pour les courbes) consiste à construire, à partir d'une surface donnée X, une **surface minimale** X' par "contractions" successives de "courbes exceptionnelles de première espèce", qui sont celles qui apparaissent par "éclatement" d'un point (un tel éclatement remplace le point par la droite projective E constituée des directions tangentes à X en ce point. Une telle courbe E est dite "exceptionnelle de première espèce" ou encore : "-1-courbe". Cette terminologie étant due au fait qu'une telle courbe est caractérisée par les conditions numériques : $K_X \cdot E = E \cdot E = -1$, ou encore : $K_X \cdot E < 0$ et $E \cdot E < 0$).

Ces contractions ne changent pas la structure birationnelle de la variété (décrite par son corps de fonctions méromorphes), mais seulement sa structure locale au voisinage d'un

nombre fini de courbes. Bien que la structure globale de la surface soit inchangée, sa topologie et les nombres d'intersection sont en général profondément modifiés. A de rares exceptions près, les propriétés de signe du fibré canonique disparaissent par éclatement d'un point. En dimension 3, la construction de modèles minimaux présente des difficultés nouvelles considérables et a été jusqu'à ces dernières années l'obstruction majeure à la classification.

La seconde étape consiste à classifier les surfaces "minimales", c'est-à-dire celles qui ne contiennent pas de telles courbes.

Ces modèles minimaux sont donnés par le tableau ci-dessous (à équivalence birationnelle près). Dans ce tableau, on voit apparaitre clairement deux types de classes : d'une part les deux classes essentiellement réductibles aux cas des courbes par produit ou fibration, d'autre part les trois classes représentées d'une part par le plan projectif, d'autre part par les surfaces soit Abéliennes soit $K3$ (et leurs sous-revêtements), et enfin par les surfaces dites de "type général". Rappelons que les surfaces $K3$ sont déformations des hypersurfaces quartiques de $\mathbb{P}^3(\mathbb{C})$: elles dépendent de 19 "modules" (20 en considérant toutes les déformations, pas forcément projectives). Elles ont été ainsi baptisées par A. Weil "en l'honneur de Kummer, Kodaira, et de la belle montagne $K2$ au Cachemire".

Ces trois classes sont les représentants de dimension complexe 2 de notre trichotomie fondamentale : le plan projectif pour la géométrie sphérique, les surfaces Abéliennes ou $K3$ pour la géométrie "plate", et le "type général" pour l'aspect hyperbolique.

$\kappa(X)$	X	K_X	p_m	p_{12}
$-\infty$	$\mathbb{P}_2(\mathbb{C})$ plan projectif	< 0	$= 0$ $(\forall m > 0)$	$p_{12} = 0$
	$\mathbb{P}_1(\mathbb{C}) \times \mathbb{C}$, avec $g(C) \geq 1$ (surfaces réglées)			
0	surfaces abéliennes ou biellitpiques	$\equiv 0$	$p_{12m} = 1$ $(\forall m)$	$p_{12} = 0$
	$K3$ ou Enriques			
1	"Proprement elliptiques" : $X \to C$ la fibre générique est une courbe elliptique		$\sim Cm$	$p_{12} \geq 2$ $K_X^2 = 0$
2	"Type général" :	> 0	$\sim ((K_X^2)/2)m^2$	$p_{12} \geq 2$ $K_X^2 > 0$

Figure 3 : Les surfaces minimales

De manière analogue au cas des courbes, les surfaces de type général sont assez bien décrites par deux invariants topologiques : $K_X^2 = c_1^2$ et c_2 (la caractéristique topologique d'Euler-Poincaré de X), qui jouent le rôle du genre pour les courbes, et sont applés les "nombres de Chern" de X. Tout comme pour les courbes, ces deux invariants étant fixés, il n'y a qu'un nombre fini de familles de déformations de surfaces ayant ces invariants. On connait assez bien la répartition des valeurs possibles (la "géographie des nombres de Chern") de ce couple d'invariants (inégalités de Noether et de Bogomolov-Miyaoka-Yau). Par contre, ces familles ne semblent pas jouir de propriétés communes, et leur topologie ne varie pas selon un principe général connu en fonction du couple (c_1^2, c_2).

Le tableau est analogue à celui des courbes et montre aussi que le comportement asymptotique des plurigenres, tant pour les courbes que pour les surfaces, décrit également ces

classes (à l'exception des 2 premières, qu'il ne distingue pas), et peut donc tout aussi bien être utilisé comme base de la classification.

C'est justement l'observation ouvrant notre **deuxième période**, définie par l'introduction du programme d'Iitaka-Moishezon.

11 Le programme d'Iitaka-Moishezon

En 1970, la situation était celle décrite par B. Moishezon dans son court article (fort injustement oublié) aux Actes du Congrès International des Mathématiciens à Nice : "La classification est encore restreinte à celle des surfaces, et il y a peu de différence entre nos connaissances et celles des classiques de l'école Italienne."

Il poursuit en définissant ce qu'il appelle la "dimension canonique" de X (de dimension n arbitraire), (baptisée par S. Iitaka "dimension de Kodaira" et notée $\kappa = \kappa(X)$. Cette terminologie est inattendue, puisque Kodaira n'a jamais utilisé cette notion, mais seulement les nombres de Chern pour les surfaces minimales).

L'invariant (birationnel, par le théorème d'Hartogs) $\kappa(X)$ prend l'une des $(n+2)$ valeurs $\{-\infty, 0, 1, \ldots, n\}$ et est défini comme suit en fonction du comportement asymptotique possible des plurigenres $p_m = p_m(X)$:

$\kappa(X) = -\infty$ si $p_m = 0, \forall m > 0$;

$\kappa(X) = d \geq 0$ si $p_m = C_m m^d$, avec $C_m = C_m(X)$ admettant une limite finie strictement positive quand m, suffisamment divisible, tend vers $+\infty$. (On montre que c'est toujours le cas).

De plus, si $\kappa(X) \geq 0$, les systèmes linéaires $|K_X^{\otimes m}|$ (m grand et divisible), fournissent une fibration $h : X \to Y$ (appelée aujourd'hui "fibration d'Iitaka") sur une variété Y de dimension $\kappa(X)$, et dont la fibre générale F est telle que $\kappa(F) = 0$. (Cette fibration est l'application (seulement méromorphe, en général), qui associe à un point générique $x \in X$ le point de l'espace projectif $\mathbb{P}(H^0(X, K_X^{\otimes m})^*)$ défini par l'hyperplan constitué des sections s'annulant en x. On a désigné par V^* le dual complexe de l'espace vectoriel V).

Cette fibration h réduit donc dans une certaine mesure la compréhension des X arbitraires à celle des trois types suivants :

$\kappa(X) = n$, que Moishezon qualifie de "type général" (le terme est resté),

$\kappa(X) = 0$, et enfin :

$\kappa(X) = -\infty$.

Pour la première fois (explicitement du moins, mais il était très fortement suggéré dans la classification de I. Shafarevitch et ses collaborateurs), un schéma de classification par réduction de dimension (aujourd'hui appelé "programme d'Iitaka") paraissait qui soit valable en toute dimension.

Précisons les liens unissant les trois classes $\kappa = n, 0, -\infty$ avec la trichotomie :

Pour les courbes, il y a coincidence. Pour les surfaces, les notions de signe métriques et numériques coincident, et on a une très bonne concordance (birationelle) entre les classes $\kappa = 2$ et $K_X > 0$ d'une part, et les classes $\kappa = 0$ et $K_X \equiv 0$ d'autre part.

Par contre les classes $\kappa(X) = -\infty$ et X rationnelle (*i.e.* birationnelle à $\mathbb{P}_2(\mathbb{C})$) sont, comme la classification de Kodaira-Enriques le montre, très différentes : on a $\kappa(X) = -\infty$ si et seulement si la surface X est birationnelle à $C \times \mathbb{P}_1(\mathbb{C})$, où C est une courbe arbitraire.

On verra dans la troisième période la signification géométrique (conjecturale) de $\kappa(X) = -\infty$.

La même période voit la publication par S.Iitaka d'un article décrivant le même invariant et la même fibration h.

Dans un second article, S. Iitaka formule sa célèbre conjecture "$C_{n,m}$" sur l'additivité des dimensions de Kodaira : $\kappa(X) \geq \kappa(B) + \kappa(F)$ si $f : X \to B$ est une fibration (*i.e.* une application holomorphe surjective) de fibre générale F et de base B.

Cette conjecture joue un rôle central dans la classification, et a été l'un des objectifs majeurs des travaux de géométrie algébrique durant la décennie qui nous occupe (1970-80). Sa résolution dans le cas où B est de type général ou X de dimension 3 par Y.Kawamata et E.Viehweg, à l'aide de difficiles résultats de nature transcendante dûs à P. Griffiths et W. Schmid, est l'un des résultats centraux de cette période. Par la suite, la démonstration a pu être ramenée à des résultats classique de théorie de Hodge (la fermeture des formes différentielles holomorphes à pôles logarithmiques, due á P.Deligne).

Deux autres résultats de cette décennie, relatifs à la géométrie "sphérique" en dimension 3, doivent être mentionnés :

1. La solution (négative) en 1971-72 du problème de Lüroth par Iskovskih-Manin, Clemens-Griffiths et Artin-Mumford (indépendamment, et par trois méthodes et exemples très différents). En particulier, les (hypersurfaces lisses) cubiques et quartiques de $\mathbb{P}^4(\mathbb{C})$ sont des variétés de Fano non rationnelles ; les cubiques et certaines quartiques sont unirationnelles, c'est-à-dire dominées par $\mathbb{P}^3(\mathbb{C})$.
2. La classification (par Iskovskih et Shokurov) des variétés de Fano X de dimension 3 dont le second nombre de Betti $b_2 = b_2(X)$ est égal à 1. (Les variétés de Fano sont celles dont le fibré anticanonique est ample. Les espaces projectifs sont des exemples typiques de variétés de Fano). En dimension 1 et 2, seules $\mathbb{P}_1(\mathbb{C})$ et $\mathbb{P}_2(\mathbb{C})$ ont ces propriétés. Mais la classification d'Iskovskih-Shokurov fait apparaitre 17 familles de déformation de telles variétés.

Ces deux résultats montrent que la géométrie "sphérique" est en dimension 3 incomparablement plus riche et complexe qu'en dimension 1 ou 2, et revèle des aspects insoupçonnables en ces petites dimensions.

12 1980- 2000 : La dimension trois, les courbes rationnelles et le programme des modèles minimaux

Les variétés de Fano (ou unirationnelles) ont $\kappa = -\infty$. Plus généralement, il est facile de voir que les variétés **uniréglées** ont aussi $\kappa = -\infty$. On dit que X est uniréglée si elle est recouverte par des **courbes rationnelles** (ce sont celles qui sont images d'une application holomorphe $f : \mathbb{P}_1(\mathbb{C}) \to X$).

On voit que pour courbes et surfaces, la réciproque est vraie (par les tableaux de classification). Ce qui motive la :

Conjecture (*) : Si $\kappa(X) = -\infty$, alors X est uniréglée.

La difficulté de cette conjecture est que s'il est facile de voir que la présence de beaucoup de courbes rationnelles est une obstruction à l'existence de formes pluricanoniques, il est bien plus difficile de construire des courbes rationnelles (les plus rares de toutes) en se basant sur l'absence de formes pluricanoniques. On remarquera qu'il s'agit bien de traduire en termes géométriques une propriété de negativité du fibré canonique.

Or, en 1979, S. Mori a découvert une méthode révolutionnaire de construction de courbes rationnelles (en toute dimension n) dans sa solution de la conjecture d'Hartshorne : "les espaces projectifs complexes sont les seules variétés ayant un fibré tangent ample".

Cette découverte a provoqué un bouleversement considérable de la théorie de la classification, que nous allons brièvement décrire.

La production de courbes rationnelles :

Le point crucial de la démonstration de S. Mori est la construction d'UNE courbe rationnelle (sous l'hypothèse d'amplitude du fibré tangent, mais dans cette étape la seule hypothèse utilisée est l'existence d'une courbe C telle que $K_X \cdot C < 0$).

Sa construction (simple dans son principe, délicate dans sa technique) consiste à faire dégénérer ("à module constant", et en gardant un point fixé) une courbe donnée C de X en une réunion de 2 courbes, dont l'une, disons C', est rationnelle. (Ce type de dégénérescence de courbes pseudo-holomorphes est appelé en géométrie symplectique : "formation de bulles").

Une condition suffisante pour qu'une telle dégénérescence soit possible est, d'après la théorie des déformations, que le nombre d'intersection $K_X \cdot C$ soit négatif (cet entier est l'opposé de la première classe de Chern de X évaluée sur la classe d'homologie $[C]$ de C, vue comme 2-cycle entier sur X ; la négativité précédente signifie intuitivement que le fibré tangent a des directions "positives" le long de C, dans lesquelles on peut en effet espérer déplacer C, en fixant l'un de ses points. Une telle direction positive correspond (en plus faible) à l'existence d'un champ de vecteurs au premier ordre le long de la courbe, pointant dans la direction positive donnée et nul au point fixé ; l'existence d'un tel champ fournissant bien sûr les déformations souhaitées de C).

Ceci fonctionne si le corps de base est de caractéristique $p > 0$ (grâce à une utilisation décisive de l'automorphisme de Frobenius, qui (itéré) opère le miracle (penser à la formule d'Hurwitz reliant les genres de deux courbes dont l'une est revêtement, ramifié ou pas, de l'autre) de fournir des revêtements de courbes de degrés arbitrairement grand sans changer le genre. Cet argument algébrique n'a pu être remplacé à ce jour par un argument géométrique ou analytique, argument qui semble nécéssiter des idées radicalement nouvelles).

En caractéristique 0, on doit d'abord "réduire modulo p", puis "relever" C' en caractéristique 0. On peut même choisir la courbe rationnelle C' de telle sorte que sa "longueur" $\mathcal{L}(C') := -K_X \cdot C'$ soit comprise entre 1 et $n+1$ (ces bornes étant optimales par des exemples immédiats).

Le résultat de Mori peut finalement être ainsi formulé : s'il existe une courbe $C' \subset X$ telle que $K_X \cdot C' < 0$), alors X contient une courbe **rationnelle** C telle que : $1 \leq -K_X \cdot C \leq n+1$. (On peut imposer à C de rencontrer C').

Il s'agit d'une illustration spectaculaire de la "définition" de la géométrie algébrique donnée plus haut : traduire en termes géométriques les propriétés de positivité ou négativité du fibré canonique. (Mais cette "définition" est partiellement suggérée par les travaux de Mori !).

On voit donc ainsi apparaitre l'alternative suivante :

1. Ou bien $K_X \cdot C \geq 0$ pour toute courbe C tracée sur X. (On dit que K_X est "nef" (pour "numériquement effectif")). Dans ce cas, le résultat de Mori ne nous dit rien. Observons cependant que si X est une surface, ceci signifie que X est "minimale" et de dimension de Kodaira positive ou nulle (par la classification). La terminologie a été étendue : on dit alors (en toute dimension) que X est un "modèle minimal" si K_X est nef. On verra ci-dessous (voir : "la conjecture d'Abondance") ce que l'on attend (et sait en dimension 3) des modèles minimaux.
2. Ou bien K_X n'est pas nef. Alors le résultat précédent de Mori fournit une information sur la structure de X. En fait, cette analyse a été poussée beaucoup plus loin, et s'est développée en une théorie entièrement nouvelle (même pour les surfaces !) : la théorie des rayons extrêmaux et contractions extrêmales, que nous allons très grossièrement décrire.

Les rayons extrêmaux :

Dans un second article (1981), S. Mori a étendu sa méthode pour montrer que le cône des courbes effectives de X est localement polyédral dans sa partie K_X-négative, ceci en toute dimension. (Le cône des courbes effectives est le cône fermé engendré dans l'espace vectoriel réel $H_2(X, \mathbb{R})$ de l'homologie réelle de degré 2 de X par les classes des courbes (complexes) tracées dans X. La partie K_X-négative est le demi-espace constitué des classes Z telles que $K_X \cdot Z < 0$ si $c_1(X) \neq 0$, et est vide sinon. Donc : K_X n'est pas nef signifie justement que la partie K_X-négative du cône des courbes est non vide. Cette propriété de finitude provient de la considération des rayons extrêmaux de ce cône, dont S. Mori montre qu'ils sont engendrés dans la partie K_X-négative par des classes de courbes rationnelles.

Même dans le cas pourtant classique et très étudié des surfaces, cette structure était totalement insoupçonnée.

Notons que la structure du cône fermé des courbes effectives est, en général très compliquée à l'extérieur de la partie K_X-négative : les rayons extrêmaux peuvent s'accumuler près de l'hyperplan K_X-nul (un tel exemple est donné par le plan projectif éclaté en les 9 points d'intersection de deux courbes elliptiques).

Le pas suivant en ce qui concerne la complexité de ce cône est celui des surfaces $K3$ ou des variétés de Calabi-Yau de dimension 3, cas étudié entre autres par P.M.H. Wilson : la partie non cubique de la frontière du cône est encore définie par des rayons associés à des contractions birationnelles. (Le théorème de Nakai-Moishezon décrit en toute généralité ce cône fermé par une famille (en général infinie dénombrable) d'inégalités polynomiales de degrés au plus $n = \dim(X)$). Un cas simple est celui d'une variété abélienne générique : ce cône est alors défini par la positivité du nombre d'auto-intersection (défini par dualité) et du nombre d'intersection avec un fibré ample arbitraire. Il est donc l'une des composantes connexes du lieu défini par une seule inégalité polynomiale de degré n.

Les contractions extrêmales :

Toujours dans ce même article, S. Mori montre (en dimension 3, et pour X lisse) que chaque rayon extrêmal K_X-négatif R détermine une **contraction extrêmale**, qui est un morphisme (*i.e.* une application holomorphe) $\phi : X \to Y$, avec Y projective, qui contracte sur des points exactement les courbes C dont la classe d'homologie appartient au rayon R.

S. Mori classifie complètement toutes les contractions extrêmales possibles (en dimension

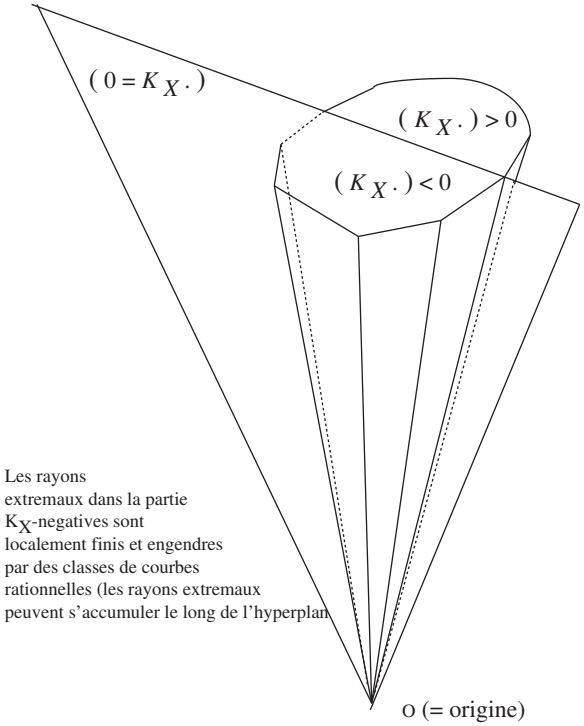


FIG. 3 – Les rayons extrêmaux : le cône fermé des courbes effectives

3, rappelons-le) en 8 types.

En dimension 2 et X lisse, la démonstration est beaucoup plus simple, et la liste similaire, mais beaucoup plus courte : il y a 3 types (classiques, mais qui n'apparaissaient pas sous cet angle) de contractions extrêmales (classifiées par le minimum $\mathcal{L}$ des "longueurs" au sens ci-dessus de toutes les courbes rationnelles dont la classe d'homologie est dans $\mathbb{R}$) :

1. $\mathcal{L} = 1$; $\phi : X \to Y$ est la contraction d'une courbe exceptionnelle de première espèce.
2. $\mathcal{L} = 2$; $\phi : X \to Y$ est un"réglage" de X (*i.e.* une fibration localement triviale de fibre $\mathbb{P}_1(\mathbb{C})$ sur une courbe Y).
3. $\mathcal{L} = 3$; $\phi : X \to Y$ est constante et $X \cong \mathbb{P}_2(\mathbb{C})$.

La première contraction de cette liste est dite birationnelle, les deux dernières "fibrantes" (car leurs fibres sont toutes de dimension strictement positive). Observons que la première simplifie la structure locale de X près du lieu exceptionnel, tandis que les deux dernières décrivent la structure globale de X très précisément.

En dimension 3, il y a 5 types de contractions birationnelles, dont 3 types ont une base Y singulière, et deux types de contractions fibrantes non constantes : les fibrations en surfaces de Del Pezzo et les fibrations en coniques. Lorsque ϕ est constante, X est donc l'une des variétés de Fano avec $b_2 = 1$, classifiées par Iskovskih-Shokurov.

L'une des premières applications de la théorie de Mori est l'étude des variétés de Fano

(qui sont précisément celles dont le cône des courbes est entièrement K_X-négatif, et dont la théorie de Mori contrôle le mieux la géométrie).

Les variétés de Fano :

Lorsque X est de Fano, le cône des courbes effectives de X est donc entièrement K_X-négatif (ceci est en fait équivalent à l'amplitude de K_X), et ce cône est donc polyédral fini.

Sa classification des contractions extrêmales permet à S. Mori en collaboration avec S. Mukai de classifier les variétés de Fano de dimension 3 avec $b_2 \geq 2$, complétant ainsi celle de Iskovskih-Shokurov. Il y a cette fois 87 familles de déformation de telles variétés, et donc 104 familles de déformation de variétés de Fano en dimension 3, au total. (Une famille "oubliée" a été récemment ajoutée a la liste initiale, par ces mêmes auteurs. Espérons que la liste actuelle est maintenant complète !).

La méthode de Mori de construction de courbes rationnelles a permis aussi d'obtenir en toute dimension des résultats qualitatifs fondamentaux concernant les variétés de Fano : celles-ci sont **rationnellement connexes**, c'est-à-dire que 2 points arbitraires peuvent y être joints par une courbe rationnelle (tout comme dans un espace projectif, où 2 points peuvent être joints par une droite). (On notera qu'il s'agit encore de traduire géométriquement une propriété de négativité du fibré canonique : lorsque cette négativité est maximale, on veut avoir le plus de courbes rationnelles possible).

Les variétés rationnellement connexes (et donc aussi les variétés de Fano) sont simplement connexes. Enfin, se basant sur leur connexité rationnelle Kollár-Miyaoka-Mori ont réussi à démontrer que, pour tout n, il n'y a qu'un nombre fini de familles de déformation de variétés de Fano de dimension n.

Leur argument repose essentiellement sur un résultat intermédiaire d'un très grand intérêt : le lemme de "collage" (de courbes rationnelles). Ce lemme permet de déformer en une courbe irréductible C' la réunion d'une courbe irréductible arbitraire C tracée sur X et d'un nombre suffisant (dépendant de C et X) de courbes rationnelles "libres" R_i de X qui rencontrent C. Ici, "libre" signifie que le fibré tangent à X le long des R_i n'a pas de direction négative, ou encore que l'on peut déplacer chacune des R_i dans toutes les directions ("librement", donc). Lorsque C est rationnelle, C' l'est aussi.

L'une des conséquences de ce lemme est que si X est rationnellement connexe, tout ensemble fini de X est contenu dans une courbe rationnelle (irréductible) de X. (Exercice : démontrer ceci pour X un espace projectif. On pourra remarquer qu'une courbe rationnelle est dans ce cas paramétrée par $n+1$ polynômes homogènes en 2 variables, tous de même degré).

On notera que les variétés rationnellement connexes sont donc celles réalisant le maximum de l'abondance en courbes rationnelles, qualitativement.

Enfin, toujours grâce à l'utilisation des courbes rationnelles, Cho-Miyaoka-Shepherd-Baron ont récemment montré que $\mathbb{P}_n(\mathbb{C})$ est la seule variété de dimension n ayant une contraction de longueur $n+1$. Cette caractérisation d'aspect technique est d'une extrême utilité dans les applications : elle signifie concrètement que $\mathbb{P}_n(\mathbb{C})$ est la seule variété de dimension n dont deux points arbitraires peuvent être joints par une "droite", une droite étant par définition une courbe rationnelle dont les déformations ne dégénèrent pas en plusieurs composantes (contrairement à une conique du plan, qui dégénère en une réunion de deux

droites, par exemple).

La géométrie des courbes rationnelles fournit par ailleurs une approche à la classfication des variétés différentiables d'holonomie spéciale, par la théorie des twisteurs.

Les modèles minimaux en dimension trois :

La liste des contractions extrêmales en dimension 3 montre que, contrairement au cas des surfaces, celles-ci ne peuvent être itérées telles quelles, puisque certaines produisent des singularités. La méthode de construction de Mori ne semble pas pouvoir être étendue au cas où X est singulière, car elle repose sur la théorie des déformations de courbes dans une variété non singulière. On est ainsi confronté à deux problèmes :

1. Définir une classe de singularités admissibles, auxquelles on puisse espérer étendre la théorie (des rayons extrêmaux et contractions extrêmales).
2. Etendre cette théorie dans le cadre des singularités admissibles.

La définition des singularités admissibles (dites "terminales") a été proposée dès 1983 par M. Reid, suite à ses travaux sur les singularités des modèles canoniques : ce sont celles qui "préservent l'adjonction" de manière comparable aux variétés lisses (les formes pluricanoniques d'une désingularisation doivent s'annuler sur chaque composante du lieu exceptionnel).

Dans ce cadre, Y. Kawamata a réussi à construire en toute dimension les contractions extrêmales des rayons extrêmaux (et des faces) K_X-négatifs du cône des courbes. L'approche de Y. Kawamata est cohomologique, et repose sur une utilisation très sophistiquée d'une généralisation importante du théorème d'annulation de Kodaira, qui lui est due (ainsi qu'à E. Viehweg). Néammoins, la méthode de Mori, basée sur le recours à la caractéristique $p > 0$, reste à ce jour la seule permettant de voir que ces rayons extrêmaux sont engendrés par des classes de courbes rationnelles.

Les "flips" :

Malheureusement, la partie la plus difficile du programme n'était pas encore réalisée (nous sommes en 1986) : dans la succession des contractions extrêmales, on peut voir surgir de "petites" contractions (birationnelles, dont le lieu exceptionnel est de codimension 2 ou plus. Ceci ne peut se produire en dimension 3 que si X est singulière. En dimension 3, ce lieu exceptionnel est donc une courbe. En dimension 2, cette situation n'existe pas). Ces contractions produisent des singularités qui ne sont pas admissibles. A nouveau, M. Reid a indiqué la voie à suivre en dimension 3 : ces courbes devraient être "flippées", c'est-à-dire qu'un processus de chirurgie holomorphe devrait leur être appliqué, remplaçant la courbe C^- , qui est K_X-négative, par une courbe C^+, qui soit $K_{X'}$-positive.

La construction de ces flips, très difficile, a finalement été réalisée par S. Mori en 1988, dans un troisième article. Elle est basée sur des travaux antérieurs de S. Mori, Y. Kawamata, V. Shokurov, M. Reid et J. Kollár, et sur un monumental travail de classification des voisinages analytiques des courbes C^-. Cette construction permet enfin la construction de modèles minimaux en dimension 3 ; plus précisément : partant de X de dimension 3, on aboutit, après un nombre fini de contractions extrêmales et de "flips", à une variété X' avec singularités admissibles, birationnelle à X, et on a l'alternative suivante :

1. Ou bien X' est un modèle minimal ($K_{X'}$ est "nef")
2. Ou bien X' a une contraction extrêmale fibrante (en particulier : X', et donc X, est uniréglée, car la fibre générale d'une contraction extrêmale fibrante est une variété de

Fano).
Ceci achève la construction des modèles minimaux en dimension 3.

Pour obtenir une classification des variétés de dimension 3 analogue à celle de Kodaira-Enriques, il restait encore à établir :

La Conjecture d'Abondance : Si $K_{X'}$ est "nef", alors $K_{X'}^{\otimes m}$ est engendré par ses sections globales pour m assez grand et divisible (il est "abondant" en sections. Cette propriété n'est pas vraie pour tous les fibrés en droites, même en supposant leur nombre d'intersection avec toute courbe tracée sur x strictement positive).

Remarquons que la réciproque est immédiate : si $K_X^{\otimes m}$ est engendré par ses sections globales pour un $m > 0$, alors K_X est nef. Remarquons aussi que cette conjecture n'est pas de nature numérique, et est donc très subtile : par exemple, si $K_X \equiv 0$, il s'agit de montrer que K_X est un élément de torsion dans le"groupe de Picard" de X, qui est un groupe de Lie complexe abélien à composante neutre compacte.

La première étape pour établir la conjecture est certainement de montrer que ce fibré a des sections non nulles. Elle a été accomplie par Y. Miyaoka en 1988 : si X' est un modèle minimal (de dimension 3), alors $\kappa(X') \geq 0$.

Ce résultat établit en même temps la difficile conjecture (*) en dimension 3 : si $\kappa(X) = -\infty$, alors X est uniréglée.

Pour établir ce résultat, Y. Miyaoka a raffiné la méthode de Mori, l'appliquant au fibré cotangent au lieu du fibré canonique. Il a ainsi obtenu (en toute dimension) un remarquable critère d'"uniréglage" (inconnu même pour les surfaces) : si la restriction du fibré tangent à X à une courbe assez générale (dans un sens précis) de X a un sous-fibré de degré positif, alors X est uniréglée.

Le reste de la conjecture (obtenue par les efforts conjoints de Y. Miyaoka et Y. Kawamata) consiste (comme pour les surfaces, mais les difficultés techniques sont d'un tout autre ordre de grandeur) en l'analyse des déformations infinitésimales des diviseurs pluricanoniques ainsi produits.

Le résultat final est une classification comparable a celle de Kodaira-Enriques pour les surfaces, réalisée en une dizaine d'années !

Notons cependant trois différences fondamentales avec la classification de Kodaira-Enriques : l'absence de bornes effectives sur les plurigenres pour déterminer la dimension de Kodaira, l'absence d'une géographie des nombres de Chern pour les modèles minimaux en dimension trois, et enfin l'absence de classification des variétés de dimension 3 avec $\kappa = 0$ (cette dernière est due a l'abondance de telles variétés : on connait plusieurs centaines de familles de variétés de dimension 3 de Calabi-Yau (simplement connexes et à fibré canonique trivial, elles jouent à première vue le rôle des surfaces $K3$ dans la classification, mais celles-ci sont "symplectiques" au sens holomorphe, ce qui les fait appartenir à une autre classe, beaucoup plus restreinte, dans la décomposition de Bogomolov).

On ne dispose à l'heure actuelle d'aucun principe de classification, même conjectural, pour les variétés de Calabi-Yau de dimension 3. Une grande partie des recherches de ces 10 dernières années ad'ailleurs été consacrée à l'étude de la géométrie énumérative des courbes rationnelles sur ces variétés. Cette direction a bien sûr été privilégiée à la suite des "prédictions" de géométrie énumérative faites (en 1990) par certains spécialistes de la théorie des champs (conformes, supersymétriques). A l'origine cette prédiction concernait

le nombre N_d de courbes rationnelles de degré fixé d, arbitraire, sur une quintique lisse X générique de l'espace projectif de dimension quatre (X est alors bien une variété de Calabi-Yau de dimension 3). On savait classiquement que $N_1 = 2875$; $N_2 = 609\,250$ avait été obtenu en 1986 ; $N_3 = 317\,206\,375$ n'a été (laborieusement) confirmé qu'en 1995.

Aujourd'hui, ces "prédictions" issues de la physique sont partiellement confirmées par des arguments mathématiques pour les quintiques de dimension 3 : leurs invariants de Gromov-Witten ont étés (définis et) calculés. Mais l'interprétation de ces invariants comme nombres de courbes rationnelles affectées de multiplicités adéquates est encore très ouverte : on ne sait pas si la quintique générique contient un nombre fini de courbes rationnelles de degré $d > 11$ fixé, par exemple.

Nous renvoyons à [66] et [26] pour ces développements.

Remarques :

1. La dimension 4 ne semble pas accessible avec les mêmes méthodes : la classification des petites contractions sur laquelle repose la construction des flips ne semble pas réalisable en dimension 4.
2. La théorie des modèles minimaux a été étendue au cas des"log-modèles minimaux" par V. Shokurov. Il ne considère plus K_X mais $K_X + D$, où D est un diviseur effectif (à coefficients rationnels). Cette extension n'augmente pas la difficulté technique, et offre beaucoup plus de flexibilité. Il a rédigé un preprint [28] contenant une construction des "flips" logarithmiques en dimension 4 par d'autres techniques, plus algébriques (le problème étant de montrer que certaines algèbres sont de type fini), et une récurrence sur la dimension.
3. Faute de place ou de compétence, de nombreux résultats en relation directe avec les sujets traités ici n'ont pu être abordés.

Un exemple est le traitement des théorèmes d'annulation par la théorie des idéaux multiplicateurs de A. Nadel, et des applications des méthodes transcendantes (estimées L^2 de Hörmander, théorie des courants, travaux de P. Lelong et H. Skoda) développées entre autres par J.P. Demailly et Y.T. Siu.

D'autres exemples sont l'étude des variétés de Calabi-Yau, la symétrie miroir, la cohomologie quantique qui nécessiteraient un développement séparé.

Un autre développement important obtenu récemment par Abramovich-Karu-Matsuki-Włodarczyk est le théorème de factorisation faible, qui décompose toute application birationnelle entre deux variétés projectives lisses (en dimension arbitraire) en une composée d'éclatements de centres lisses et de leurs inverses. Voir [24] pour un exposé et des références complémentaires.

D'autre part, nombre de résultats obtenus trop récemment pour en estimer l'impact n'ont pu être inclus.

Citons néammoins :

a. L'invariance des plurigenres par déformation (projective) de la structure complexe, par Y.T. Siu.
b. L'existence d'une section pour une fibration de base une courbe et de fibre générique rationnellement connexe, par Graber-Harris-Starr. (Si la base est rationnelle, l'espace total est donc rationnellement connexe).
c. L'introduction de la classe des variétés "spéciales", qui combinent les géométries "sphé-

rique" et "plate", et fournissent une dichotomie "spéciale"-"type général", exprimée (pour X arbitraire) par une unique fibration fonctorielle $c : X \to Y$ à fibres spéciales et base de type général, dans un sens "orbifolde" approprié.

Cette fibration permet (conjecturalement, sous réserve de version "orbifolde" de la conjecture $C_{n,m}$) de décomposer toute variété projective en ses composantes plates, sphériques et hyperboliques.

Elle semble aussi fournir une description qualitative simple des propriétés arithmétiques, hyperboliques, topologiques de X. (Voir [12]).

13 Remerciements

Je remercie beaucoup P.Y. Gaillard pour ses critiques très constructives et sa patience. Elles ont permis d'améliorer considérablement ce texte, à la fois dans sa forme et dans son contenu. Je remercie aussi vivement L. Bonavero pour son aide dans la réalisation des tableaux et dessins, ainsi que pour les nombreuses corrections et améliorations qu'il a apportées au texte original.

La bibliographie est très restreinte, et vise plutôt à permettre de retrouver une bibliographie plus complète dans celle des ouvrages cités.

Les articles originaux ne sont pas toujours cités. De nombreux développements importants n'ont malheureusement pu être inclus. On a tenté de fournir des références facilement accessibles et lisibles, se concentrant dans un premier temps moins sur les techniques que sur les idées directrices.

Bibliographie

[**Remarque :**] Cette bibliographie est très restreinte, et vise plutôt à permettre de retrouver une bibliographie plus complète dans celle des ouvrages cités. Les articles originaux ne sont pas toujours cités. De nombreux développements importants n'ont malheureusement pu être inclus. On a tenté de fournir des références facilement accessibles et lisibles, se concentrant dans un premier temps moins sur les techniques que sur les idées directrices.

Avant 1970

[1] J. Dieudonné, *Cours de géométrie Algébrique*, PUF, Collection Sup., 1974.
Le premier volume est de nature historique. Bibliographie très complète sur les textes classiques.

[2] C. Houzel, *Fonctions elliptiques et intégrales abéliennes* in *Abrégé d'histoire des Mathématiques*, (sous la direction de J.Dieudonné), Hermann, 1-113, 1978.
Remarquable description très détaillée des développements de cette théorie durant la période 1669-1900.

[3] O. Zariski, *Algebraic Surfaces*, 2e édition, Springer Verlag, Erg. der Math. 61., Springer Verlag, 1971.

Surfaces

[4] K. Kodaira, *On compact complex analytic surfaces I,II,III*, Ann. Math. 71, 1960 ; 77, 1963 ; 78, 1963.

[5] I. Shafarevitch et al., *Algebraic Surfaces*, Proc. Stekhl. Inst. 75, 1965.

[6] A. Beauville, *Surfaces algébriques complexes*, Astérisque 54, SMF, 1978.
Exposé concis , démonstrations claires par une approche classique (*i.e.* : pré-Mori), de brèves mais intéressantes indications historiques.

[7] W. Barth, C. Peters, A. Van de Ven, *Compact complex surfaces*, Erg. der Math. 4 (3. Folge), Springer-Verlag, 1984. (La seconde édition comportera une classification du point de vue des contractions extrêmales).

[8] *Surfaces K3, modules et périodes*, Séminaire Palaiseau, Astérisque 126, SMF, 1985.

Programme d'Iitaka-Moishezon

[9] S. Iitaka., *On D-dimension of Algebraic Varieties*, J. Math. Soc. Japan 23, 1971.

[10] B. Moishezon, *Algebraic Varieties and compact complex spaces*, Actes du C.I.M. Nice, Vol.2, pp.643-48, 1970.

[11] K. Ueno, *Classification of Algebraic Varieties and Compact Complex Spaces*, LNM 439, Springer Verlag, 1975.

[12] F. Campana, *Special varieties and classification theory*, Preprint 2001, (math.AG/0110051).

Variétés de Fano

[13] V. Iskovskih,*Anticanonical models of three-dimensional algebraic Varieties*, J. Soviet Math. 13, 1980.

[14] V. Shokurov, *The existence of a straight line on Fano threefolds* Math. USSR Izv. 26, 1175-209, 1979.

[15] S. Mori, S. Mukai, *Classification of Fano threefolds with* $b_2 \geq 2$, Man. Math. 36, 1981.

[16] F. Campana, *Connexité rationnelle des variétés de Fano*, Ann. Sc. ENS. 25, 1992.

[17] J. Kollár, S. Mori, Y. Miyaoka, *Rational connecteness and boundedness of Fano Manifolds*, J. Diff. Geom. 36, 1992.

[18] O. Debarre., *Variétés de Fano*, Exposé Bourbaki, 1999.

[19] V. Iskovskih, Y. Prokhorov, *Fano Varieties*, Algebraic Geometry V, EMS 47, Springer Verlag, 1999.

Travaux de Mori

[20] S. Mori, *Projective manifolds with ample tangent bundle*, Ann. Math. 110, 1979.

[21] S. Mori, *Threefolds whose canonical bundle is not numerically effective*, Ann. Math. 116, 1981.

[22] S. Mori, *Flip theorem and the existence of minimal models for threefolds*, J. AMS. 1, 1988.

Contractions extrêmales.

[23] Y. Kawamata, K. Matsuki, K. Matsuda, *Introduction to the Minimal model problem*, In Av. Stud. Pure Math. 10, Kinikuniya-North-Holland, 1987.

On y trouvera un exposé de l'approche cohomologique de Kawamata des contractions extrêmales.

[24] L. Bonavero, *Factorisation faible des applications birationnelles (d'après Abramovich, Karu, Matsuki, Włodarczyk et Morelli)*, Séminaire Bourbaki, Exposé 880, Novembre 2000.

Variétés de Calabi-Yau, symétrie Miroir, invariants de Gromov-Witten

[25] C. Voisin, *Symétrie miroir*, Panoramas et synthèses 2, SMF 1996.

Concis, mais technique.

[26] D. Cox, S. Katz, *Mirror Symmetry and Algebraic Geometry*, Math. Surveys and Monographs 58. AMS 1998.

Excellente introduction à tous les aspects de cette théorie encore en gestation. Très complet et (néanmoins) lisible, avec de bonnes descriptions des motivations physiques, et aussi les applications à la géométrie énumerative.
Pour une introduction agréable aux invariants de Gromov-Witten du point de vue symplectique, on pourra consulter :

[27] M. Audin, *Invariants en géométrie symplectique, via les courbes holomorphes*, Panoramas et synthèses 11, SMF 1-59, 2001.

"Log-flips"

[28] V. Shohurov, *PL-flips*, Preprint 1-175, Aout 2000.

Ouvrages d'exposition

On y trouvera beaucoup d'autres références, en particulier les travaux de Reid et Shokurov.

[29] H. Clemens, J. Kollár, S. Mori, *Higher Dimensional complex geometry*, Astérisque 166, 1988.

[30] J. Kollár et al., *Flips and Abundance for Algebraic Threefolds*, Astérisque 211, SMF, 1992.

[31] J. Kollár, S. Mori, *Birational geometry of Algebraic Varieties*, Cambridge Tracts in Math. 134, Cambridge University Press, 1998.

[32] O. Debarre, *Higher Dimensional Algebraic Geometry*, Universitext. Springer Verlag, 2001.

Exposition très claire et concise, avec démonstrations et un minimum de prérequis, des sujets décrits ici.

Géométries en dimension trois : de H. Seifert à W. Thurston

M.Boileau[1]
Laboratoire Émile Picard,
CNRS UMR 5580, Université Paul Sabatier,
118 Route de Narbonne,
F-31062 TOULOUSE Cedex 4,
France

1 Introduction

Dans le cycle des conférences internationales de mathématiques consacrées à *Quelques questions de Géométrie et Topologie*, organisées en 1935 par l'Université de Genève, W. Threlfall [69] commence sa conférence de la façon suivante :

> *"Nous savons tous que le problème d'homéomorphie des variétés à n dimensions, posé par Poincaré , est un des plus intéressants et des plus importants de la Géométrie. Il est intéressant en soi car il a donné naissance à la Topologie combinatoire ou algébrique, théorie comparable par son importance à la Théorie des fonctions classique. Il est important par ses applications à la Cosmologie, où il s'agit de déterminer l'aspect de l'espace de notre intuition et de la physique. Dès l'instant où le physicien a envisagé la possibilité de considérer l'espace de notre intuition, espace où nous vivons, comme clos, la tâche du mathématicien est de lui proposer un choix d'espaces clos, et même de les énumerer tous, comme il le ferait pour les polyèdres réguliers. . . .*
>
> *Malheureusement le problème n'est complètement résolu que pour deux dimensions."*

Plus de soixante ans après, ce problème n'est toujours pas résolu en dimension 3, bien que d'énormes progrés dans la compréhension de la géométrie et de la topologie des variétés tridimensionnelles aient été accomplis.

[1]boileau@picard.ups-tlse.fr

Comme le souligne W. Threlfall dans sa conférence, un résultat majeur de la fin du 19ième siècle et du début du 20ième siècle est la preuve du Théorème d'uniformisation des surfaces de Riemann. Ce théorème, énoncé par F. Klein [30] en 1883, a été démontré indépendemment par H. Poincaré [43] et par P. Koebe [33] en 1907.

Théorèrme d'uniformisation de Riemann
Toute surface fermée admet une métrique riemannienne à courbure constante $K \equiv +1, 0$ *ou* -1. *Elle s'obtient donc comme quotient de la sphere* $\mathbb{S}^2$, *du plan Euclidien* $\mathbb{E}^2$ *ou du plan hyperbolique* $\mathbb{H}^2$ *par l'action libre et discontinue d'un groupe d'isométries.*

La formule de Gauss -Bonnet montre qu'une même surface ne peut admettre deux structures modelées sur des géométries différentes. De plus la géométrie détermine le type topologique de la surface.

On ne peut pas espérer avoir un résultat analogue en dimension 3. Threlfall [69] remarque que la variété produit $S^1 \times \mathbb{S}^2$ ne peut pas admettre une métrique riemannienne à courbure sectionnelle constante, car son revêtement universel $\mathbb{S}^2 \times \mathbb{R}$ a deux bouts, contrairement à $\mathbb{S}^3$, à $\mathbb{E}^3$ et $\mathbb{H}^3$ qui n'ont qu'un seul bout.

Ce n'est que quarante ans plus tard, avec les travaux fondamentaux de W. Thurston [**?**], qu'une théorie géométrique des variétés en dimension 3 a vraiment émergé. Cette théorie est parfaitement résumée par sa conjecture qui domine depuis bientôt trente ans la topologie en dimension 3 :

Conjecture de Thurston : L'intérieur d'une *variété compacte, orientable, de dimension 3, admet une décomposition le long d'une collection finie (peut-ête vide) de sphères* $\mathbb{S}^2$ *et de tores* $\mathbb{T}^2$ *plongés, essentiels, disjoints et deux à deux non paralleles, en sous-variétés dont l'intérieur (aprés avoir bouché les composantes de bord sphériques par des boules) porte une métrique riemannienne homogène et complète.*

La conjecture de Thurston concerne toutes les variétés tridimensionnelles, dont elle a relançé l'étude d'un point de vue géométrique. Elle englobe en particulier la conjecture de Poincaré :

Conjecture de Poincaré : *Toute variété tridimensionnelle fermée et simplement connexe est homéomorphe à la sphère standard* $\mathbb{S}^3$.

Cette conjecture formulée par Poincaré en 1904 [42] est un des grands problèmes ouverts en topologie. Elle a guidé beaucoup des développements de la topologie en petite dimension.

Dés 1977, W. Thurston a annoncé un théorème d'hyperbolisation qui est une contribution fondamentale à sa conjecture et qui a boulversé l'étude des variétés de dimension 3. En particulier, il a mis en lumière la place prépondérante de la géométrie hyperbolique dans cette dimension.

Le but de cet article est de présenter la conjecture de Thurston et son théorème d'hyperbolisation dans leur contexte historique, ainsi que certaines des idées qui sont apparues en dimension trois à cette occasion. Pour simplifier l'exposition nous supposerons toutes les variétés orientables.

1.1 Les huit géométries homogènes en dimension trois

Une variété riemannienne X est dite *homogène* si son groupe d'isométries $Isom(X)$ agit transitivement. On appelle *géométrie* une variété riemannienne complète X, simplement connexe et homogène dont le groupe d'isométries $Isom(X)$ est maximal : c'est-à-dire qu'il n'existe aucune métrique riemannienne sur X invariante par $Isom(X)$ dont le groupe d'isométries contient strictement $Isom(X)$. Deux géométries X et X' sont *équivalentes* s'il existe un difféomorphisme $\phi : X \to X'$ qui conjugue les actions de $Isom(X)$ et $Isom(X')$.

Soient X une géométrie et $\Gamma \subset Isom(X)$ un sous-groupe discret, agissant proprement et librement sur X. L'espace quotient X/Γ est une variété lisse, munie d'une métrique riemannienne homogène qui est localement isométrique à celle de X. Son groupe fondamental est isomorphe à Γ.

Une variété lisse M (éventuellement munie d'un bord) admet une *X-structure* si son intérieur est difféomorphe à un quotient X/Γ comme ci-dessus. Une variété est dite *géométrique* si elle admet une X-structure pour une géométrie X.

Une géométrie X est *isotrope* si la courbure sectionnelle de X est constante. Pour chaque dimension $n \geq 2$ il y a seulement trois géométries isotropes, à équivalence prés : la *géométrie elliptique* $\mathbb{S}^n$, la *géométrie euclidienne* $\mathbb{E}^n$ et la *géométrie hyperbolique* $\mathbb{H}^n$, à courbure sectionnelle constante respectivement égale à $+1, 0, -1$.

La notion de géométrie trouve son origine dans le mémoire présenté en 1854 par B. Riemann [57] dans lequel il introduit la notion de variété à courbure constante et énonce la classification des géométries correspondantes. A la suite des travaux de Clifford [**?**], puis de F. Klein [31], W. Killing [27] posa en 1891 le problème de déterminer toutes les variétés riemanniennes fermées à courbure sectionnelle constante de dimension n (qu'il désigne sous le nom d'*espaces de Clifford-Klein*). Il montra qu'une telle variété peut-être obtenue comme le quotient de l'espace elliptique $\mathbb{S}^n$, euclidien $\mathbb{E}^n$ ou hyperbolique $\mathbb{H}^n$ par l'action libre et discontinue d'un groupe d'isométries, réduisant ainsi le problème à l'étude des actions libres et co-compactes des sous-groupes d'isométries de ces géométries.

La notion de X-structure a été introduite en 1931 par O. Veblen et J. H. C. Whitehead [80], à la suite de l'article de E. Cartan [6] sur les variétés riemanniennes localement homogènes.

En dimension 2 la situation est trés spéciale en ce sens qu'une géométrie est toujours isotrope ; les trois géométries isotropes, modelées sur l'espace elliptique $\mathbb{S}^2$, euclidien $\mathbb{E}^2$ ou hyperbolique $\mathbb{H}^2$ sont donc les seules.

En dimension 3 nous ne nous intéressons qu'aux géométries X^3 pour laquelle il existe au moins une variété compacte tridimensionnelle dont l'intérieur porte une X^3-structure de volume fini. Nous appelons ces géométries les *géométries utiles*.

Les trois géométries isotropes sont utiles, mais il existe aussi des géométries *anisotropes* qui sont utiles ; par exemple les géométries produits $\mathbb{S}^2 \times \mathbb{E}^1$ et $\mathbb{H}^2 \times \mathbb{E}^1$, qui ont respectivement pour quotient les variétés fermées $\mathbb{S}^1 \times \mathbb{S}^2$ et $\mathbb{S}^1 \times F_g^2$, où F_g^2 est une surface hyperbolique, fermée, orientable de genre $g \geq 2$.

Thurston a observé que d'autres variétés fermées, fibrées en cercles $\mathbb{S}^1$ où en tores $\mathbb{T}^2$, admettent une géométrie anisotrope mais qui n'est pas un produit. En fait, à coté des trois géométries isotropes et des deux géométries produit, il n'existe que trois autres géométries

utiles en dimension 3, correspondant aux groupes de Lie réels Nil, $\widetilde{SL}(2,\mathbb{R})$ et Sol (cf. [77] et aussi [62]).

Voici la liste des huit géométries X^3 utiles en dimension 3, suivant la taille du stabilisateur $G_x \subset Isom^+X$ d'un point $x \in X^3$.

Les 3 géométries isotropes pour lesquelles $G_x = SO(3)$

(1) La Géométrie elliptique $\mathbb{S}^3$

Toute variété elliptique est le quotient de la sphère $\mathbb{S}^3$ par un groupe fini de transformations orthogonales.

La classification des variétés elliptiques tridimensionnelles a été entamée par Hopf [24] en 1926 et complètement achevée par H. Seifert et W. Threlfall entre 1930 et 1932 dans deux articles [69]. Ces variétés sont toutes orientables. Une des clefs de la classification est l'observation que tout sous-groupe fini de transformation orthogonales, agissant librement sur la sphère $\mathbb{S}^3$, commute avec l'action d'un sous-groupe orthogonal à un paramètre. Les variétés elliptiques tridimensionnelles admettent donc toutes une partition en cercles telle que chaque cercle admet un voisinage tubulaire saturé (i.e. réunion de cercles).

Ce phénomène a conduit Seifert à étudier dès 1933 la classe des variétés tridimensionnelles admettant ce type de fibration en cercles, que l'on appelle depuis *fibration de Seifert*. Nous reviendrons sur ces variétés au paragraphe suivant, car elles ont joué un rôle important dans l'étude des variétés de dimension 3.

(2) La géométrie Euclidienne $\mathbb{E}^3$

D'aprés le théorème de Bieberbach ([6] 1911), toute variété euclidienne fermée est le quotient du tore $\mathbb{T}^3$ par un groupe fini de transformations euclidiennes. La classification des variétés Euclidiennes tridimensionnelles a été obtenue en 1934 par W. Nowaski [49] en déterminant dans la liste des 17 groupes critallographiques de dimension 3 ceux qui agissent librement sur l'espace euclidien $\mathbb{E}^3$. Une preuve plus conceptuelle a été donnée indépendamment en 1935 par W. Hantzsche et H. Wendt [23]. Il y a 6 variétés euclidiennes fermées, orientables, de dimension 3.

(3) La géométrie hyperbolique $\mathbb{H}^3$

C'est la géométrie de dimension 3 la plus compliquée et la plus riche. Un théorème de Poincaré permet de construire beaucoup d'exemples de variétés hyperboliques par identification des faces d'un polyèdre hyperbolique.

Le premier exemple d'une variété tridimensionnelle, hyperbolique, complète et à volume fini a été construit par Gieseking en 1912 dans sa thèse [20], en identifiant deux à deux les faces d'un tétraèdre idéal, régulier. Cet exemple a été un peu oublié par la suite. C'est une variété non-orientable dont le revêtement d'orientation est l'extérieur du noeud à quatre croisements. La structure hyperbolique complète sur l'extérieur du noeud à quatre croisements a été redécouverte de façon complètement différente en 1975 par Riley [58]. Par contre, ce n'est qu'en 1931 que Löbell a construit le premier exemple d'une variété hyperbolique fermée, en recollant le long de leurs faces huit copies d'un polyèdre hyperbolique à angles droits, ayant 14 faces [34]. C'est en 1933 qu'est apparu l'espace dodécaédral hyperbolique de Seifert et Weber, obtenu en identifiant les faces d'un dodécaèdre régulier hyperbolique,

d'angles dièdres $2\pi/5$ [86]. D'autres exemples proviennent de la géométrie arithmétique. Mais il a fallu attendre les travaux de Thurston (1977) et son théorème d'hyperbolisation pour avoir un critère topologique d'existence de structures hyperboliques sans avoir à exhiber explicitement celles-ci.

Les 4 géométries anisotropes pour lesquelles $G_x = SO(2)$

Dans ce cas X^3 est un fibré en cercles ou en droites au dessus de l'espace elliptique $\mathbb{S}^2$, euclidien $\mathbb{E}^2$ ou hyperbolique $\mathbb{H}^2$.

(4) La géométrie produit $\mathbb{S}^2 \times \mathbb{E}^1$

$\mathbb{S}^2 \times \mathbb{S}^1$ et $\mathbb{R}P^3 \sharp \mathbb{R}P^3$ sont les seules variétés orientables ayant cette géométrie.

(5) La géométrie produit $\mathbb{H}^2 \times \mathbb{E}^1$

Une variété admettant cette géométrie est fibrée au sens de Seifert et a pour revêtement fini un produit $F_g \times \mathbb{S}^1$, $g > 1$.

(6) La géométrie Nil

Elle correspond au groupe de Lie nilpotent d'Heisenberg, formé des matrices triangulaires supérieures de la forme :

$$\{\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} : \quad x, y, z \in \mathbb{R}\}$$

Une variété admettant cette géométrie a pour revêtement fini un fibré en cercles $\mathbb{S}^1$ sur le tore $\mathbb{T}^2$ ayant une classe d'Euler non nulle.

La suite exacte $\mathbb{R} \to Nil \to \mathbb{R}^2$ montre qu'une telle variété est fibrée au sens de Seifert.

(7) La géométrie $\widetilde{SL}(2, \mathbb{R})$

Elle correspondant au revêtement universel $\widetilde{SL}(2, \mathbb{R})$ du groupe de Lie $SL(2, \mathbb{R})$. C'est aussi le revêtement universel du fibré unitaire tangent $\widetilde{T_1}\mathbb{H}^2$ du plan hyperbolique $\mathbb{H}^2$. Une variété admettant cette géométrie est fibrée au sens de Seifert et a pour revêtement fini un fibré en cercles, de classe d'Euler non nulle, sur une surface hyperbolique, fermée et orientable. Ces variétés sont toutes fermées et ne sont jamais des fibrés en surfaces.

La géométrie anisotrope pour laquelle $G_x = \{id\}$

(8) C'est la géométrie correspondant au groupe de Lie résoluble Sol

Le groupe de Lie Sol correspond à l'extension scindée : $\mathbb{R}^2 \to Sol \to \mathbb{R}$ où l'action de $t \in \mathbb{R}$ sur $\mathbb{R}^2$ est donnée par la matrice

$$\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}.$$

Une variété orientable admettant cette géométrie a pour revêtement à deux feuilles un fibrée en tores, avec une monodromie $\phi : \mathbb{T}^2 \to \mathbb{T}^2$ qui est Anosov, i.e. ayant deux valeurs propres réelles.

La classification des géométries utiles en dimension 3 peut alors s'énoncer de la façon suivante :

Classification des géométrie de dimension 3[71], [61] :*Une géométrie utile en dimension 3 est équivalente à l'une des huit géométries décrites ci-dessus. De plus il n'existe pas deux géométries équivalentes dans cette liste.*

La description des huit géométries utiles en dimension 3 montrent qu'une variété fermée peut admettre au plus une seule de ces géométries.

1.2 Des variétés fibrées de Seifert aux variétés graphées de Waldhausen

A la suite de ses travaux avec W. Threlfall, H. Seifert a étudié dans son article de 1933 [64] les variétés qui admettent une fibration en cercles généralisée, appelée *fibration de Seifert.*

Une *variété de Seifert* est une variété compacte orientable partitionnée en courbes simples fermées, appelées *fibres* de telle sorte que toute fibre admette pour voisinage tubulaire un tore solide fibré, obtenu de la façon suivante : on considère un cylindre circulaire et solide, fibré par des segments parallèles à son axe, et on identifie les deux bases après une rotation d'angle $2\pi\beta/\alpha$ de l'une par rapport à l'autre. Les entiers α et β sont premiers entre eux et satisfont la condition $0 \leq \beta < \alpha$. Si $\alpha > 1$ la fibre centrale, provenant de l'axe du cylindre, est dite *exceptionnelle*, d'invariants la fraction $\beta/\alpha \in (\mathbb{Q}/\mathbb{Z})^\star$. Sinon elle est dite *régulière.* Il s'agit d'une généralisation de la notion de fibrés en cercles, en ce sens que l'on permet l'existence de fibres exceptionnelles, autour desquelles les fibres voisines s'enroulent plusieurs fois.

Seifert [64] a décrit un système fini d'invariants qui permet de classer ces variétés fibrées, à homéomorphisme conservant les fibres près. Ce ne sont pas toujours des invariants topologiques : il existe des variétés pouvant admettre des fibrations de Seifert non équivalentes, comme par exemple la sphère $\mathbb{S}^3$ ou les espaces lenticulaires. Il faudra attendre les travaux de E. J. Brody [11] en 1960, de F. Waldhausen [81] en 1967 et de P. Orlik, E. Vogt et H. Zieschang [**?**] en 1968 pour obtenir la classification topologique complète des variétés de Seifert.

Dans son article original Seifert montre que la seule variété de Seifert simplement connexe est la sphère $\mathbb{S}^3$. Il montre aussi que le revêtement universel d'une variété de Seifert est homéomorphe à $\mathbb{S}^3$, $\mathbb{R}^3$ ou $\mathbb{S}^2 \times \mathbb{R}$.

Cependant, il semble que ce n'est que quarante ans plus tard, avec l'intervention des idées de Thurston, qu'on a réalisé que le revêtement universel d'une variété de Seifert peut-être muni d'une métrique riemannienne homogène pour laquelle son groupe fondamental agit par isométries. Il s'agit de l'une des six géométries homogènes, qui ne correspond ni à la géométrie hyperbolique, ni à la géométrie Sol. Ainsi toute variété de Seifert est géométrique.

Les variétés de Seifert, bien qu'à priori très spéciales, vont jouer un rôle important dans l'étude topologique des variétés de dimension 3, et en particulier dans la décomposition

canonique d'une variété irréductible le long d'une famille finie de tores plongés et essentiels. Nous commençons par présenter cette décomposition dans le cas particulier des variétés dites *graphées*.

Les variétés graphées tridimensionnelles apparaissent naturellement en géométrie analytique dans l'étude des singularités isolées de courbes planes et de surfaces.

L'intersection transverse d'une sphère S^3_ε, centrée à l'origine et de rayon $\varepsilon > 0$ assez petit, avec une courbe algébrique plane, ayant une singularité isolée à l'origine, est un entrelacs $L^1 \subset S^3_\varepsilon$ dit *torique itéré* : son extérieur $S^3_\varepsilon \backslash \mathcal{N}(L^1)$ est obtenu en recollant le long de leurs bords respectifs des variétes admettant des fibrations de Seifert, qui ne coïncident pas sur les tores de recollements (cf. [10]).

Dans le cadre des variétés fermées, les structure graphées ont été utilisées dès 1961 par D. Mumford [48] pour étudier la topologie des singularités normales de surfaces algébriques. La résolution abstraite d'une telle singularité fait apparaître un diviseur exceptionnel, à croisements normaux, dont un voisinage tubulaire s'obtient comme *un plombage* de fibrés en disques $\mathbb{D}^2$ sur des surfaces orientables trouées. C'est une variété compacte, orientable de dimension 4 à *coins* dont le bord est une variété fermée de dimension 3 qui porte une structure dite *graphée*. Cette structure correspond à une collection finie $\mathcal{T}$ de tores plongés, disjoints, qui découpent la variété en sous-variétés admettant une fibration de Seifert. En général ces fibrations de Seifert ne se recollent pas le long de certains tores de la famille $\mathcal{T}$.

Dans le langage géométrique introduit par Thurston, une variété graphée admet toujours une décomposition en morceaux géométrique ne faisant jamais intervenir la géométrie hyperbolique.

La classification topologique des variétés graphées a été commencée par R. von Rondow en 1962 [**?**], à l'instigation de F. Hirzebruch. Mais c'est réellement F. Waldhausen dans ses deux articles fondamentaux en 1967 [81] qui a introduit la notion générale de *variété graphée* et qui en a complètement décortiqué la topologie. En particulier, il a résolu pour ces variété le problème d'homéomorphie. Il a aussi montré que ces variétes vérifient la conjecture de Poincaré , et plus généralement la conjecture du revêtement universel.

Une variété orientable de dimension 3 est *irréductible* si toute sphère $\mathbb{S}^2$ plongée borde une boule dans cette variété.

La conjecture suivante implique la conjecture de Poincaré :

Conjecture du revêtement universel :*Le revêtement universel d'une variété fermée, orientable et irréductible est homéomorphe à* $\mathbb{S}^3$ *ou* $\mathbb{R}^3$.

Cette conjecture a été motivée par les travaux de C. Papakriakopoulos [52, 2] dans le milieu des années 1950. Il a donné une preuve du *Lemme de Dehn*, publié en 1910 dans un article de M. Dehn [16] mais avec une preuve erronée. Il a démontré aussi deux théorèmes fameux, le *Théorème du Lacet* et le *Théorème de la sphère*, qui font un lien entre les propriétés homotopiques et topologiques d'une variété de dimension 3 (cf. aussi [65]). Ces théorèmes sont à la base de presque tous les résultats obtenus depuis sur les variétés tridimensionnelles.

Le Théorème de la sphère montre en particulier que le revêtement universel d'une variété orientable, irréductible et de groupe fondamental infini, est contractile. Il est naturel dans

ce cas de se demander quel est son type topologique et s'il est toujours homéomorphe à l'espace $\mathbb{R}^3$.

A la fin des années 1960, F. Waldhausen [83] a obtenu une série de résultats remarquables, en particulier sur cette question, pour une vaste classe de variétés englobant la pluspart des variétés graphées. Il s'agit des *variétés hakeniennes*, qui font l'objet du prochain paragraphe.

1.3 Les variétés hakeniennes

Une surface F^2, proprement plongée dans une variété irréductible M^3 est *essentielle* si le groupe fondamental de chaque composante connexe de F^2 s'injecte dans celui de M^3 et qu'aucune composante connexe de F^2 ne borde une boule ou ne peut être déformée dans le bord ∂M.

Une variété est *hakenienne* si elle est compacte, irréductible et qu'elle contient une surface essentielle. Ces variétés ont été étudiées par W. Haken au début des années 1960 [21, 2]. F. Waldhausen à la fin des années 60 [83] a établi leurs propriétés fondamentales et a montré la place centrale qu'elles occupent dans l'étude des variétés tridimensionnelles.

Il existe beaucoup d'exemples importants de variétés hakeniennes : toute variété irréductible ayant un premier groupe d'homologie infini est hakenienne. En particulier toute variété irréductible à bord non vide est hakenienne. Un exemple typique est l'extérieur d'un nœud dans $\mathbb{S}^3$.

Il existe des variétés fermées hyperboliques qui ne sont pas hakeniennes, comme l'espace dodécaédral de Seifert et Weber [86]. Cependant ce n'est qu'à la suite des travaux de Thurston ([70], Ch.4) qu'on a su construire beaucoup d'exemples de variétés hyperboliques non hakeniennes.

L'existence d'une surface essentielle F^2, proprement plongée dans M^3, permet de découper la variété le long de cette surface et d'obtenir ainsi une nouvelle variété M^3_F qui est irréductible. Si M^3_F n'est pas une réunion finie de boules, elle est hakenienne et le processus peut être itéré. Le théorème de finitude de Haken montre qu'après un nombre fini de pas ce processus donne une réunion finie de boules : une telle suite finie de variétés, obtenues par découpages successifs le long de surfaces essentielles, est appelée une *hiérarchie* de M^3. L'existence d'une hiérarchie s'est révélée fondamentale dans la preuve de beaucoup de résultats importants concernant les variétés hakeniennes.

F. Waldhausen a pu montrer ainsi que ces variétés vérifient la conjecture du revêtement. Il a aussi démontré le théorème de rigidité topologique suivant, qui permet de résoudre le problème d'homéomorphie pour ces variétés :

Théorème de rigidité topologique[83] :*Une équivalence d'homotopie entre deux variétés hakeniennes fermées peut-être déformée en un homéomorphisme.*

Dans le cas où l'une des variétés fibre sur le cercle, ce théorème avait été démontré dès 1962 par J. Stallings [66]. C'est jusqu'à maintenant le seul résultat général de rigidité topologique, connu en dimension trois.

F. Waldhausen avait observé qu'une variétés graphée, irréductible et de groupe fondamental infini, a un revêtement fini qui est une variété hakenienne. Cette observation l'a conduit a faire la conjecture suivante [85] :

Conjecture du revêtement (Waldhausen) : *Toute variété fermée, orientable, irréductible et de groupe fondamental infini, admet un revêtement fini qui est une variété hakenienne.*

On espère en fait qu'il existe toujours un revêtement fini ayant un premier groupe d'homologie infini. Cette conjecture reste largement ouverte, même dans le cas d'une variété hyperbolique.

1.4 De Kneser à Jaco, Shalen et Johannson

Le cadre topologique de la conjecture de Thurston est donné par un découpage de la variété compacte, orientable M^3, le long d'une famille finie de sphères $\mathbb{S}^2$ et de tores $\mathbb{T}^2$ plongés de façon disjointe et essentielle dans M^3, en morceaux canoniques. L'existence de cette décomposition est un résultat central pour l'étude topologique des variétés de dimension 3. Elle est l'aboutissement de travaux importants s'étendant sur presque 50 années, à partir des travaux de à H. Kneser en 1929 [32].

Cette décomposition s'effectue en deux étapes :

La décomposition suivant des sphères essentielles $\mathbb{S}^2$

La première étape correspond à la décomposition par somme connexe d'une variété tridimensionnelle en variétés premières. L'existence d'une telle décomposition finie est due à H. Kneser [32]. La preuve de l'unicité des facteurs, dans le cas des variétés orientables, est esquissée dans l'article de Kneser ; une preuve complète est due à J. Milnor [32].

Théorème de décomposition en facteurs premiers([25], [32]) :*Toute variété tridimensionnelle, compacte, orientable et sans composante sphérique dans le bord, est homéomorphe à la somme connexe d'un nombre fini de variétés irréductibles et de copies de $\mathbb{S}^1 \times \mathbb{S}^2$. De plus les facteurs irréductibles sont uniques à homéomorphismes préservant l'orientation près.*

On a vu que les facteurs $\mathbb{S}^1 \times \mathbb{S}^2$ sont géométriques. Le théorème de Kneser permet donc de restreindre la conjecture de Thurston au cas des variétés irréductibles.

La décomposition suivant des tores essentiels $\mathbb{T}^2$

Il s'agit d'une décomposition plus subtile de chaque facteur irréductible de M^3, suivant une famille maximale de tores *canoniques*, plongés dans chaque facteur.

Un tore $\mathbb{T}^2$ plongé dans une variété compacte, orientable M^3 est dit *canonique* s'il est essentiel et s'il peut être disjoint de tout autre tore plongé essentielle par une *isotopie* (i.e. une déformation formée par une famille à 1 paramêtre d'homéomorphismes)

Théorème de décomposition torique([26], [27]) :*Toute variété hakenienne, compacte, orientable, contient une collection finie (peut-être vide) maximale de tores plongés, canoniques, disjoints et deux à deux non parallèles. Une telle famille de tores canoniques est unique à isotopie prés et découpe la variété en sous-variétés atoroïdales ou admettant une fibration de Seifert.*

Une variété M est *atoroïdale* si tout sous-groupe abélien de rang 2 du groupe fondamental de M est conjugué à un sous-groupe contenu dans le groupe fondamental d'une composante du bord.

Le *théorème du tore* montre qu'une variété hakenienne est atoroïdale si et seulement si elle n'est pas de Seifert et ne contient aucun tore plongé essentiel. Il a été annoncé en 1970 par F. Waldhausen [84], mais la preuve n'est apparue qu'en 1972 [18] (cf. [62]).

C'est H. Schubert en 1953 [34] qui le premier a étudié de façon systématique les tores plongés essentiels dans l'extérieur d'un noeud dans la sphère $\mathbb{S}^3$.

Le théorème de décomposition torique pour une variété hakenienne a été démontrée en 1975 par W. Jaco et P. Shalen [26], et indépendamment par K. Johannson [27]. C'est là une étape cruciale dans la compréhension des variétés tridimensionnelles.

A la même époque, A. Marden [36] a mis en évidence l'importance des variétés hakeniennes et des résultats de Waldhausen, dans l'étude des structures hyperboliques géométriquement finies en dimension trois. Il a reformulé, en terme de variétés de dimension trois, le critère de Klein-Maskit [37] de recollement de structures hyperboliques géométriquement finies et il a suggéré l'utilisation des hiérarchies pour leur étude. C'est probablement lui qui, le premier, a posé le problème de caractériser toplogiquement une variété hyperbolique ([36], p.461) ; il a remarqué qu'une telle variété doit être irréductible et que son groupe fondamental ne peut pas avoir de centre.

Les travaux de Waldhausen [82, 22, 4], de Jaco-Shalen [26] et Johannson [27], en mettant en évidence l'importance des variétés atoroïdales dans l'étude des variétés de dimension 3, ont soulevé la question de savoir si ces variétés admettent une structure particulière. Ils ont montré aussi que le groupe fondamental d'une variété atoroïdale possède certaines des propriétés algébriques des sous-groupes discrets du groupe des isométries de l'espace hyperbolique $\mathbb{H}^3$.

C'est sans doute ce qui a conduit W. Thurston à proposer un cadre géométrique pour l'étude des variétés de dimension trois. Dès 1977 il a annoncé que l' intérieur d'un morceau atoroïdal, à bord non vide et non homéomorphe à un I-fibré sur le tore, porte une métrique hyperbolique. C'est un cas particulier de son théorème d'hyperbolisation pour les variétés hakeniennes. Nous présentons maintenant ce résultat.

1.5 Le Théorème d'hyperbolisation de Thurston

En 1977, W. Thurston a annoncé que la conjecture de géométrisation est vraie pour les variétés hakeniennes. La preuve découle d'un théorème d'hyperbolisation pour ces variétés, dont voici l'énoncé général :

Théorème d'hyperbolisation[72] :*Soit M une variété tridimensionnelle, compacte et orientable. On suppose que la variété M est irréductible, hakenienne et n'est pas homéomorphe à un fibré en intervalles sur la bouteille de Klein. L'intérieur de la variété M admet une structure hyperbolique complète si et seulement si la variété M est atoroïdale.*

Les deux conditions, irréductibilité et atoroïdalité, sont nécessaires pour que l'intérieur de M admette une structure hyperbolique complète. C'est une conséquence directe de

l'irréductibilité du revêtement universel $\mathbb{H}^3$ et de la classification des sous-groupes abéliens d'un sous-groupe, de type fini, discret et sans torsion de $PSL_2(\mathbb{C})$.

Par contre, la condition d'être hakenienne n'est pas du tout nécessaire, comme le montre l'exemple de l'espace dodécaédral hyperbolique de Seifert et Weber [86]. On conjecture qu'elle peut être remplacée par la condition (cette fois-ci nécessaire) que le groupe fondamental de M est infini.

Le théorème d'hyperbolisation s'applique en particulier à l'extérieur d'un nœud non trivial dans S^3 qui n'est ni un nœud torique, ni un nœud satellite. R. Riley [58], aprés avoir explicitement construit des exemples de structures hyperboliques sur des complémentaires de nœuds et entrelacs dans $\mathbb{S}^3$, avait d'ailleurs conjecturé le théorème d'hyperbolisation pour ces variétes dés 1973.

Le théorème de chirurgie hyperbolique de Thurston montre comment construire énormément de variétés fermées, hyperboliques qui ne sont pas hakeniennes. Partant d'une variété irréductible M, ne contenant aucune surface fermée essentielle et à bord un tore incompressible, Thurston déforme la structure hyperbolique complète (donnée par le théorème d'hyperbolisation) en une structure hyperbolique non complète, mais dont le complété est topologiquement une variété fermée, obtenue en collant un tore solide le long du bord ∂M.

Théorème de chirurgie hyperbolique([70], Ch.5) : *Soit M une variété hakenienne atoroïdale et $\mathbb{T}^2 \subset \partial M$ une composante torique incompressible. A l'exception d'un nombre fini d'entre elles, toute variété, obtenue en collant un tore solide le long du tore $\mathbb{T}^2$, admet une structure hyperbolique complète.*

Ce théorème est une indication trés forte en faveur d'une réponse positive à la conjecture de géométrisation de Thurston. Il est apparu pour la première fois dans les notes de Thurston ([70], Chap. 5, (cf. [4], [8], Appendice B)).

C'est dans ses notes de cours de Princeton, des années 1978/80 [70], [71], que Thurston a commencé à exposer les techniques et les résultats sur lesquels se fondent la preuve de son théorème d'hyperbolisation. Par leur richesse et la profonde nouveauté du point de vue adopté, ces notes sur la *Géométrie et Topologie de dimension 3* ont jeté les bases d'une véritable révolution hyperbolique en dimension trois. Elles sont immédiatement devenues et demeurent le texte de référence dans ce domaine.

En 1980, Thurston a présenté un plan détaillé et précis de la preuve du théorème d'hyperbolisation dans des notes sur les conférences qu'il avaient données à Bowdoin [74] (cf. [72]). Ces conférences sont à l'origine du premier texte publié, du à John Morgan [45], donnant une vue d'ensemble de la preuve, avec une rédaction détaillée de certaines parties.

Puis Thurston a écrit la première partie de sa preuve dans [75, 8] ainsi que la preuve du cas fibré sur le cercle dans [76]. Le reste de la preuve devait alors être reconstituée à partir de [74] et des chapitres 8, 9 et 13 de ses notes [70].(cf. [45]).

En 1985, J. Morgan et P. Shalen [46, 2] ont donné une preuve différente, de nature plus algébrique, des résultats principaux de [75, 8].

En suivant une approche complètement différente de celle de Thurston, C. McMullen [39, 2] a donné en 1989 une preuve du *théorème du point fixe*, qui est le cœur de la démonstration du théorème d'hyperbolisation des variétés qui ne fibrent pas sur le cercle. Cette approche trouve son origine dans l'observation par J. Hubbard d'un lien entre la conjecture de Kra,

sur la norme de l'opérateur Théta associé au revêtement d'une surface de Riemann, et le théorème du point fixe de Thurston.

A partir des travaux de McMullen, J.P. Otal [51] a écrit en 1996 une preuve complète du théorème d'hyperbolisation des variétés qui ne fibrent pas sur le cercle. Entre-temps, il avait donné une nouvelle preuve pour le cas des variétés fibrées dans [50].

A la même époque M. Kapovich [28] a écrit lui aussi une preuve complète du théorème d'hyperbolisation pour les variétés qui ne fibrent pas sur le cercle, mais en reconstituant l'approche originale de Thurston.

Les détails complets de la preuve sont donc maintenant accessibles dans la monographie de M. Kapovich [28], et les monographies de J. P. Otal [50, 2].

La preuve du théorème d'hyperbolisation, comme beaucoup de théorèmes difficiles en dimension 3, utilise l'existence d'une hiérarchie pour la variété hakenienne. Celle-ci peut donc être construite inductivement, à partir d'une collection finie de boules, en recollant à chaque étape les morceaux le long de sous-surfaces incompressibles dans leurs bords. L'étape initiale de la récurrence est donnée par le théorème d'Andreev, qui caractérise les polyèdres pouvant être réalisés dans l'espace hyperbolique $\mathbb{H}^3$ avec un volume fini et des angles dièdres droits. Le cœur de la démonstration est alors un théorème de recollement de structures hyperboliques, auquel se réduit l'étape de récurrence. Une jolie construction de miroitage du bord de la variété permet à Thurston de ramener l'étape de récurrence à un théorème de recollement qui fait intervenir tout le bord de la variété.

Dans ce qui suit nous décrivons la preuve dans le cas plus simple où le bord de de la variété ne contient pas de tore.

On désigne par τ une involution du bord qui échange les composantes connexes par paires et renverse leur orientation : il y a une partition du bord de la variété M, $\partial M = \partial^+ M \sqcup \partial^- M$, et un difféomorphisme renversant l'orientation $f : \partial^+ M \to \partial^- M$ tel que $\tau = (f, f^{-1})$.

On note alors M/τ la variété fermée, obtenue en identifiant deux à deux les composantes connexes de ∂M par τ.

Avec ces notations, l'étape principale de la preuve du théorème d'hyperbolisation de Thurston est le théorème de recollement suivant :

Théorème de recollement : *Soit M une variété hyperbolique, à bord incompressible, non vide. Soit $\tau : \partial M \to \partial M$ une involution qui échange les composantes connexes par paires en renversant leur orientation. La variété M/τ est hyperbolique si et seulement si elle est atoroïdale.*

La preuve de ce théorème distingue deux cas suivant que la variété M admet ou non un revêtement fini qui est le produit d'un intervalle par une surface compacte. La démonstration dans les deux cas est différente.

Dans la preuve du théorème de recollement Thurston a fait le lien entre la topologie et les méthodes analytiques développées par L. Ahlfors et L. Bers [1], [2], [5] dans les années 1960 pour étudier les groupes kleiniens. Il utilise aussi les travaux plus géométriques de A. Marden [36], B. Maskit [37, 2] et D. Sullivan [62] dans les années 1970.

1.6 Le cas fibré sur le cercle

Dans ce cas, quitte à passer à un revêtement fini, on peut supposer que la variété M/τ fibre sur le cercle, avec pour fibre la surface F et pour *monodromie* $f \in Diff^+(F)$, telle que $\tau = (f, f^{-1})$. La variété est alors obtenue comme la *suspension* $F \times_f \mathbb{S}^1 := F \times [0,1]/\{(x,0) \sim (f(x),1)\}$ du difféomorphisme $f \in Diff^+(F)$.

La condition d'atoroïdalité est alors équivalente au fait que le difféomorphisme f est *pseudo-Anosov*, c'est-à-dire que son action sur l'ensemble des classes de conjugaison de lacets non triviaux dans $\pi_1(F)$ n'admet pas d'orbite périodique [73], [17].

Le cas fibré sur le cercle peut alors s'énoncer sous la forme suivante :

Théorème d'hyperbolisation pour les variétés fibrées :*Soient F une surface compacte, orientable et de caractéristique d'Euler $\chi(F) < 0$ et $f \in Diff^+(F)$ un difféomorphisme de F. La variété $F \times_f \mathbb{S}^1$, obtenue par suspension du difféomorphisme f est hyperbolique si et seulement si f est pseudo-Anosov.*

Une preuve détaillée de ce cas a été donnée dans [50] et [71] (cf. aussi [41], [62]).

W. Thurston pense que les fibrés sur le cercle occupent une place importante au sein des variétés hyperboliques, puiqu'il a fait la conjecture suivante :

Conjecture (Thurston) : *Toute variété tridimensionnelle, hyperbolique, complète et à volume fini a un revêtement fini qui fibre sur le cercle.*

A ce jour, cette conjecture est toujours ouverte.

1.7 Le cas non fibré

Dans le cas des variétés hakeniennes qui ne fibrent pas sur le cercle Thurston ramène la preuve du théorème de recollement à un théorème de point fixe.

Pour une variété hyperbolique M, à bord incompressible, on note $\mathcal{GF}(M)$ l'espace des structures hyperboliques géométriquement finies sur M. C'est l'analogue de l'espace de Teichmüller pour une surface.

On note $\mathcal{T}(\partial M)$ l'espace de Teichmuller du bord ∂M de la variété M. Soit $\partial : \mathcal{GF}(M) \to \mathcal{T}(\partial\mathcal{M})$, l'application de restriction ; d'après le théorème de finitude d'Ahlfors [1], elle fait correspondre à une structure hyperbolique géométriquement finie sur M la structure conforme d'aire finie, induite sur son bord ∂M.

Le théorème d'Ahlfors-Bers [2] montre alors que cette application $\partial : \mathcal{GF}(M) \to \mathcal{T}(\partial M)$ est une bijection.

Pour traduire le théorème de recollement en un théorème de point fixe pour une certaine application de $\mathcal{T}(\partial M)$ dans lui-même, Thurston a introduit l'*application d'épluchage* σ : $\mathcal{T}(\partial M) \to \mathcal{T}(\partial\widetilde{M})$, où $\partial\widetilde{M}$ est le bord de la variété M, munie de l'orientation renversée.

Nous définissons maintenant cette application. Dans ce qui suit l'indice F désigne la projection sur le facteur de $\mathcal{T}(\partial M)$ correspondant à l'espace de Teichmuller $\mathcal{T}(F)$ de la composante $F \subset \partial M$.

D'après le théorème d'Ahlfors-Bers, une classe de structures conformes $s \in \mathcal{T}(\partial M)$ est induite par une structure hyperbolique géométriquement finie g_s sur M. Pour cette structure hyperbolique, chaque composante connexe $F \subset \partial M$ détermine un revêtement *quasi-fuchsien*

$(\overline{M}_F, \overline{g}_s)$ de la variété hyperbolique (M, g_s) (cf. Maskit [38]). C'est une structure hyperbolique géométriquement finie sur la variété produit $F \times I$. L'application de restriction induit donc une bijection $\partial_F : \mathcal{GF}(F\times I) \to \mathcal{T}(F)\times\mathcal{T}(\widetilde{F})$ telle que $\partial_F(\overline{g}_s) = (s_F, s'_F)$, où $s_F \in \mathcal{T}(F)$ est la composante sur le facteur $\mathcal{T}(F)$ de $s \in \mathcal{T}(\partial M)$.

L'application d'épluchage $\sigma : \mathcal{T}(\partial M) \to \mathcal{T}(\partial\widetilde{M})$ est définie par l'élément $\sigma(s) \in \mathcal{T}(\partial\widetilde{M})$ dont la composante sur le facteur $\mathcal{T}(\widetilde{F})$ est $\sigma(s)_F = s'_F$.

Dans la suite on note $\tau^\star : \mathcal{T}(\partial M) \to \mathcal{T}(\partial\widetilde{M})$ l'involution induite par l'involution τ. Thurston a traduit le critère de recollement de Klein-Maskit [37, 2] en un critère d'existence de point fixe pour la composée $\tau^\star \circ \sigma : \mathcal{T}(\partial M) \to \mathcal{T}(\partial M)$.

Critère de recollement : *Soit M une variété hyperbolique, vérifiant les hypothèses du théorème de recollement. Si l'application $\tau^\star \circ \sigma : \mathcal{T}(\partial M) \to \mathcal{T}(\partial M)$ admet un point fixe, la variété quotient M/τ est hyperbolique.*

Avec ce critère, le Théorème de recollement est une conséquence du théorème suivant :

Théorème du point fixe :*Soit M une variété hyperbolique, à bord incompressible, non vide et qui n'est pas fibrée en intervalles. Soit $\tau : \partial M \to \partial M$ une involution qui échange les composantes par paires en renversant leur orientation. L'application $\tau^\star \circ \sigma : \mathcal{T}(\partial M) \to \mathcal{T}(\partial M)$ admet un point fixe si et seulement si la variété quotient M/τ est atoroïdale.*

Il existe deux preuves très différentes de ce théorème.

La première est due à Thurston. Elle utilise beaucoup d'outils géométriques qu'il a introduits à cet effet, et qui ce sont révélés fondamentaux dans l'étude des variétés hyperboliques de dimension 3. Ces notions sont présentées dans les chapitres 8 et 9 des notes de Thurston [70], qui sont les chapitres les plus difficiles à lire. C'est cette approche qu'a suivi M. Kapovich dans son livre [28].

Une autre preuve a été donnée en 1989 par C. McMullen [39, 2]. Elle suit une approche totalement différente, plus proche de la théorie de Teichmuller classique et des méthodes analytiques à la Ahlfors-Bers. Elle permet d'obtenir une majoration de la norme de la dérivée de l'application d'épluchage en un point. C'est l'approche qu'a suivi J.P. Otal dans son article [51].

1.8 Quelques résultats dans le cas non hakenien

Dans le cas non hakennien, le théorème de décomposition torique se réduit à montrer qu'une variété tridimensionnelle, fermée, orientable, irréductible et qui n'est pas atoroïdale admet une fibration de Seifert.

La preuve découle des travaux de P. Scott [62, 22], P. Tukia [79], G. Mess [42] (voir aussi [35], D. Gabai [13], A. Casson et D. Jungreis [14]), qui s'étalent sur dix ans (1980-1990). Elle s'appuie non seulement sur les travaux de Waldhausen, de Jaco-Shalen et de Johannson, mais elle fait aussi intervenir les techniques plus récentes de la géométrie des groupes. Ces travaux ont permis de donner la caractérisation homotopique suivante des variétés de Seifert :

Théorème :*Une variété tridimensionnelle, compacte, orientable, irréductible et de groupe fondamental infini, admet une fibration de Seifert si et seulement si son groupe fondamental contient un sous-groupe infini cyclique qui est distingué.*

Pour l'extérieur d'un noeud dans $\mathbb{S}^3$ ce théorème a été démontré en 1966 par G. Burde et H. Zieschang [12]. Il a été généralisé en 1967 par F. Waldhausen [82] au cas d'une variété hakenienne.

La conjecture de géométrisation de Thurston se réduit donc à deux conjectures d'uniformisation pour les variétés non hakeniennes, ne faisant plus intervenir que la géométrie elliptique et la géométrie hyperbolique. La première conjecture est une extension de la conjecture de Poincaré, englobant la conjecture de Smith sur le redressement des actions libres des sous-groupes finis de $Diff^+(\mathbb{S}^3)$.

Conjectures d'uniformisation :

Soit M une variété orientable, fermée, irréductible et non hakenienne :

1.(Poincaré-Smith) La variété M admet une métrique sphérique si et seulement si son groupe fondamental est fini.

2.(Thurston) La variété M admet une structure hyperbolique si et seulement si son groupe fondamental est infini et qu'elle est atoroïdale.

A ce jour, ces conjectures restent largement ouvertes.

En 1981 Thurston [72], [78] a annoncé la preuve de la conjecture de géométrisation pour les variétés admettant une symétrie axiale :

Théorème : *Soit M une variété de dimension 3 fermée, orientable et irréductible. Soit $f \in Diff^+(M)$ un difféomorphisme périodique et non trivial. Si f n'agit pas librement sur M, la variété M admet une décomposition géométrique équivariante par f.*

Ce théorème implique la conjecture de Smith pour les actions finies, cycliques et non libres sur la sphère S^3, dont la preuve a été un des résultats importants en topologie au tout début des années 1980 (cf. [44]).

Une preuve détaillée de ce théorème est maintenant accessible dans [8]. Cet énoncé est en fait un cas particulier d'un théorème concernant la géométrie des orbi-variétés (cf. [9], [15]). Il s'agit d'une généralisation naturelle de la notion de variété, introduite en 1956 par I. Satake [57], et dont l'importance, dans l'étude des variétés de dimension 3, a été remarquée par Thurston [70]. En particulier, Thurston a utilisé la notion d'orbivariété dans la preuve de son théorème d'hyperbolisation pour l'étape du miroitage du bord de la variété.

Bibliographie

[1] L. Ahlfors, *Finitely generated Kleinian groups*, 1964, Amer. J. Math., vol **86**, pp. 413-429

[2] L. Ahlfors, L. Bers, *Riemann's mapping theorem for variable metrics*, Ann. of Math. vol **72**, 1960, pp. 385–404

[3] E. M. Andreev, *On convex polyhedra in Lobačeskiĭ space*, 1970, Mat. USSR Sbornik, vol **10**, pp. 413-440

[4] R. Benedetti, C. Petronio, *Lectures on Hyperbolic Geometry*, ed. Universitext, Springer-Verlag, Berlin, 1992

[5] L. Bers, *Uniformization, moduli and Klenian groups*, 1972, vol **4**, Bull. London Math. Soc., pp. 257-300

[6] L. Bieberbach, *Über die Bewegungsgruppen der Euklidischen Räume I* , Math. Ann., 1911, pp. 297-336, vol **70**

[7] F. Bonahon, *Bouts des variétés hyperboliques de dimension 3*, 1986, vol **124**, Ann. of Math., pp. 71-158

[8] M. Boileau, J. Porti, *Geometrization of 3-orbifolds of cyclic type*, Astérisque, vol **272**, 2001

[9] M. Boileau, B. Leeb, J. Porti, *Uniformization of small.*, C. R. Acad. Sci. Paris Sér. I Math., vol **332**, 2001, pp. 57-62

[10] K. Brauner, *Zur Geometrie der Funktioen zweier Veränderlichen : II-IV*, Abh. Math. Sem. Hamburg, 1928, pp. 1-54, vol **6**

[11] E.J. Brody, *The topological classification of the lens spaces*, Ann. of Math., 1960, pp. 163-184, vol **71**

[12] G. Burde, H. Zieschang, *Eine Kennzeichnung der Torusknoten*, Math. Ann., 1966, pp. 169-176, vol **167**

[13] E. Cartan, *L'application des espaces de Riemann et l'analysis situs*, Assoc. Avanc. Sci. Lyon, vol **53** , 1926, (œuvres complètes, Vol. III, Gauthier- Villars, Paris (1952), 993-995

[14] A. Casson, D. Jungreis, *Convergence group and Seifert fibered 3-manifolds*, 1994,**118** , Invent. Math., pp. 441-456.

[15] D. Cooper, C. Hodgson, S. Kerchkoff, *Three dimensional Orbifolds and Cone Manifolds*, Mathematical Society of Japan Memoirs, 2000,**5**

[16] M. Dehn , *Über die Topologie des dreidimensionalen Raumes*, Math. Ann., 1910, pp. 137-168,**69**

[17] A. Fathi, F. Laudenbach, et V. Poenaru, *Travaux de Thurston sur les surfaces*, Astérisque,**66-67**, 1979

[18] C. Feustel, *On the torus theorem for closed 3-manifolds*, Trans. Amer. Math. Soc., 1976, pp. 45-57,**217**

[19] D. Gabai, *Convergence groups are Fuchsian groups*, Ann. of Math., 1992,**136**, pp. 447-510

[20] H. Giesieking, *Analytische Untersuchung über topologische Gruppen*, Ber. Thèse Univ. Münster, 1912

[21] W. Haken, *Theorie der Normalflächen*, Acta Math., 1961, pp. 245-375,**195**

[22] W. Haken, *Some results on surfaces in 3-manifolds*, Studies in Modern topology, Math. Assoc. Amer., ed. Prentice Hall, 1968, pp. 39-98

[23] W. Hantzsche, H. Wendt, *Dreidimensionale euklidische Raumformen*, Math. Ann., 1935, pp. 593-611,**110**

[24] H. Hopf , *Zum Clifford-Kleinschen Raumproblem*, Math. Ann., 1929, pp. 313-339,**95**

[25] W. H. Jaco, *Lectures on three-manifold topology*, CBMS Regional conference Series in Mathematics,**43**, American Mathematical Society, Providence, R.I., 1980

[26] W. H. Jaco, P. B. Shalen, *Seifert fibred spaces in 3-manifolds*, Mem. Amer. Math. Soc.,**220** , 1979

[27] K. Johannson, *Homotopy equivalences of 3-manifolds with boundary*, Lecture Notes in Mathematics,**761**, ed. Springer Verlag, Berlin, 1979

[28] M. Kapovich, *Hyperbolic Manifolds and Discrete Groups*, Progress in Mathematics, Birkhäuser, Boston, Inc, M.A., 2001

[29] W. Killing, *Ueber Clifford-Klein'schen Raumformen*, Math. Ann., 1891, pp. 257-278,**39**

[30] F. Klein, *Neue Beiträge zur Riemann'schen Functionentheorie*, Math. Ann., 1883, pp. 141-218,**21**

[31] F. Klein, *Zur Nicht-Euklidischen Geometrie*, Math. Ann., 1890, pp. 544-572,**37**

[32] H. Kneser, *Geschlossene Flächen in dreidimensionalen Mannigfaltigkeiten*, Jahr. Deut. Math. Verein., 1929, pp. 248-260,**21**

[33] P. Koebe, *Über die Uniformisierung beliebiger analytischer Kurven*, Nachr. Akad. Wiss. Göttingen Math.-Phys. KL., 1907, pp. 191-210

[34] F. Löbell, *Beispiele geschlossener dreidimensionaler Clifford-Kleinscher Raüme negativer Krümmung*, Ber. Verh. Sächs. Akad. Wiss. Leipzig. Math.-Phys. KL. , 1931, pp. 167-174,**83**

[35] S. Maillot, *Quasi-isometries of groups, graphs and surfaces*, 2001,**76**, Comment. Math. Helv., pp. 20-60

[36] A. Marden, *The geometry of finitely generated Kleinian groups*, 1974,**99**, Ann. of Math., pp. 383-496

[37] B. Maskit, *On Klein's combination theorem III*, 1971,**66**, ed. L. Ahlfors et al., in *Advances in the theory of Riemann Surfaces*, Ann. of Math. Studies, Princeton Univ. Press, pp. 297-316

[38] B. Maskit, *Kleinian Groups*, Springer Verlag, Berlin, 1988

[39] C. McMullen, *Iteration on Teichmüller space*, Invent. Math.,**99**, 1990, pp. 425-454

[40] C. McMullen, *Riemann surface and the geometrization of 3-manifolds*, Bull. Amer. Math. Soc., **27**, 1992, pp. 207-216

[41] C. McMullen, *Renormalization and 3-manifolds which Fiber over the Circle*, Annals of Math. Studies, Princeton University Press,**144**, 1996

[42] G. Mess, *The Seifert conjecture and groups which are coarse quasi-isometric to planes*, Preprint, 1988

[43] J.Milnor, *A unique decomposition theorem for* 3*-manifolds.*, 1962,**84** , Amer. J. Math., pp. 1-7

[44] J. W. Morgan, H. Bass, *The Smith Conjecture*, Academic Press, 1984

[45] J. Morgan, *On Thurston's uniformization theorem for three-dimensional manifolds*, 1984, ed H. Bass et J. Morgan, in *The Smith Conjecture*, Academic Press, Orlando, Fl., pp. 37-125

[46] J. Morgan, P. B. Shalen, *Valuations, trees and degenerations of hyperbolic structures, I*, 1984,**120** , Ann. of Math., pp. 401-476

[47] J. Morgan, P. B. Shalen, *Valuations, trees and degenerations of hyperbolic structures, II and III :*, 1988,**127** , Ann. of Math., pp. 403-465 and 467-519

[48] D. Mumford, *The topology of normal singularities of an algebraic surface and a criterion for simplicity.*, Inst. Hautes Études Sci. Publ. Math.,**9**, 1961, pp. 5-22

[49] W. Novacki, *Die euklidischen dreidimensionalen, geschlossen und offenen Raumformen*, Comment. Math. Helv., 1934, pp. 81-92,**7**

[50] J.-P. Otal, *Le théorème d'hyperbolisation pour les variétés fibrées de dimension 3*, Astérisque,**235**, 1996

[51] J.-P. Otal, *Thurston's hyperbolization of Haken manifolds*, Preprint, 1997

[52] C.D. Papakyriakopoulos, *On Dehn's lemma and the asphericity of knots*, Ann. of Math., 1957, pp. 1-26,**66**

[53] C.D. Papakyriakopoulos, *On Solid Tori*, Proc. London Math. Soc., 1957, pp. 281-299,**7**

[54] H. Poincaré, *Cinquième complément à l'analysis situs*, Rend. Circ. Mat. Palermo, 1904, pp. 45-110,**18**

[55] H. Poincaré, *Sur l'uniformisation des fonctions analytiques*, Acta Math., 1907, pp. 1-63,**31**

[56] J. G. Ratcliffe, *Foundations of Hyperbolic Geometry*, Graduate texts in Mathematics, ed ; Springer Verlag, Berlin, 1994

[57] B. Riemann, *Über die Hypothesen, welche der Geometrie zu Grunde liegen*, Abh. Ges. Wiss. Göttingen, 1867, pp. 133-152,**13**

[58] R. Riley, *A quadratic parabolic group*, Math. Proc. Camb. Phil. Soc., 1975, pp. 281-288,**77**

[59] I. Satake, *On a generalization of the notion of manifold*, Proc. Nat. Acad. Sci. U.S.A., 1956, pp. 359-363,**42**

[60] H. Schubert, *Knoten und Vollringe*, Acta Math., 1953, pp. 131-286,**60**

[61] P. Scott, *The geometries of 3-manifolds*, Bull. London Math. Soc., 1983, pp. 401-487,**15**

[62] P. Scott, *A new proof of the annulus and torus theorems*, Amer. J. Math., 198O, pp. 241-277,**102**

[63] P. Scott, *There are no fake Seifert fibre spaces with infinite* π_1, Ann. of Math., 1983, pp. 35-70,**117**

[64] H. Seifert, *Topologie der dreidimensionaler gefaserter Raüme*, Acta Math., 1933, pp. 147-288,**60**

[65] J. Stallings, *On the loop theorem*, Ann. of Math., 1960, pp. 12-19,**72**

[66] J. Stallings, *On fibering certain 3-manifolds*, 1962, in *Topology of 3-manifolds and related topics*, Prentice Hall, pp. 95-100

[67] D. Sullivan, *On the ergodic theory at infinity of an arbitrary discrete group of hyperbolic motions*, 1981, **97**, in *Riemann Surfaces and related topics*, Ann. of Math. Studies, Princeton Univ. Press

[68] W. Threlfall, H. Seifert, *Topologische Untersuchung der Diskontinuitätsbereiche endlicher Bewegungsgruppen des dreidimensionalen sphärischen Raumes I and II* , Math. Ann., 1930 and 1932,**104** and **107** , pp. 1-70 and 543-586

[69] W. Threlfall, *Quelques resultats récents de la topologie des variétés*, Enseign. Math., 1936,**35** , pp. 242-255

[70] W. P. Thurston, *The geometry and topology of 3-manifolds*,ed. Princeton Math. Dept., 1979

[71] W. P. Thurston, *Three-dimensional geometry and topology*, ed S. Levy, Princeton University Press,**35, 199**

[72] W. P. Thurston, *Three dimensional manifolds, Kleinian groups and hyperbolic geometry*, **6** , Bull. Amer. Math. Soc., 1982, pp. 357-381

[73] W. P. Thurston, *On the geometry and dynamics of diffeomorphisms of surfaces*,**19** , Bull. Amer. Math. Soc., 1988, pp. 417-432

[74] W. P. Thurston, *Hyperbolic structures on 3-manifolds : overall logic*, 1980, Notes of Summer Workshop at Bowdoin

[75] W. P. Thurston, *Hyperbolic Structures on 3-manifolds, I : Deformations of acylindrical manifolds* ,**124** , Annals of Math., 1986, pp. 203-246

[76] W. P. Thurston, *Hyperbolic Structures on 3-manifolds, II : Surface groups and manifolds which fiber over* $\mathbb{S}^1$, 1986, Preprint

[77] W. P. Thurston, *Hyperbolic Structures on 3-manifolds, III : Deformations of 3-manifolds with incompressible boundary* , 1986, Preprint

[78] W. P. Thurston, *Three-manifolds with symmetry*, 1982, Preprint

[79] P. Tukia, *Homeomorphic conjugates of Fuchsian groups*, J. Reine Angew. Math., 1988, pp. 1-54,**391**

[80] O. Veblen, J.H. C. Whitehead, *A set of axioms for differential geometry*, Proc. Nat. Acad. Sci. U.S.A., 1931, pp. 551-561,**17**

[81] F. Waldhausen, *Eine Klasse von 3-dimensionalen Mannigfaltigkeiten I and II*, 1967,**3** and **4** , Invent. Math., pp. 308-333 and 87-117

[82] F. Waldhausen, *Gruppen mit Zentrum und 3-dimensionale Mannigfaltig-keiten*, 1967,**6** , Topology, pp. 505-517

[83] F. Waldhausen, *On irreducible 3-manifolds which are suffiently large*, 1968,**87** , Ann. of Math., pp. 56-88

[84] F. Waldhausen, *On the determination of some 3-manifolds by their fundamental group alone*, 1969, Proc. International Symposium in Topology, Hercy-Novi,Yugoslavia, Beograd, pp. 331-332

[85] F. Waldhausen, *Some problems on* 3*-manifolds*, in *Algebraic and geometric topology*, Proc. Sympos. Pure Math., XXXII, Amer. Math. Soc., Providence, R.I. , 1978,**XXXII** , pp. 313-322

[86] C. Weber, H. Seifert, *Die beiden Dodekaederräume*, Math. Z., 1933, pp. 237-253,**37**

Autour de l'hypothèse de Poincaré

Valentin Poénaru

1 Introduction

Je vais commencer par dire, brièvement, ce que le texte qui suit contient et, aussi, ce qu'il ne contient pas.

La section suivante (II) est une brève esquisse de l'histoire de la conjecture de Poincaré en dimensions supérieures à trois ($n \geq 4$). Bien entendu cette histoire est complètement enchevêtrée avec celle de toute la topologie, au XX$^{\text{ème}}$ siècle, mais je me suis limité dans cet exposé historique aux seules choses qui touchent notre sujet (et je n'ai pas pu les aborder toutes !) j'ai commencé avec Poincaré lui-même en sautant, après, jusqu'à Hurewicz et J.H.C. Whitehead, pour finir avec Freedman et Donaldson. Que tous ceux auxquels je n'ai pas rendu assez justice m'excusent !

Les trois sections suivantes concernent la dimension $n = 3$. Il se trouve que l'hypothèse de Poincaré , question encore ouverte à ce jour, peut-être conçue, mis à part son caractère topologique évident, comme pur problème de théorie des groupes discrets, de géométrie différentielle, de calcul variationnel, d'EDP, e.a.d.s. Mis à part un très bref rappel sur le programme de Bill Thurston (section III) je n'ai rien dit sur les nombreux travaux importants concernant le sujet, dus à différents autres mathématiciens. En particulier, et je suis pleinement conscient que ceci est une grosse lacune, je n'ai rien dit sur l'approche EDP, le problème de Yamabe, le flot de Ricci, e.a.d.s.

C ertains lecteurs ne s'étonneront peut-être pas que j'ai consacré la section IV à la description de mon propre programme sur l'hypothèse de Poincaré et la section V à une importante retombée du programme en question : la preuve de la simple connexité à l'infini des revêtements universels des variétés fermées 3-dimensionnelles. La relation entre la section II et ces deux dernières sections n'est pas si ténue, puisque IV se passe (presque) exclusivement en dimension $n = 4$ et V en dimensions très grandes.

Vu le côté "réflexion philosophique" de notre livre, j'ai ajouté une dernière section V, "Mathématiques et Ontologie" qui n'est qu'un essai à caractère provisoire.

2 Brève histoire de l'hypothèse de Poincaré

Il s'agira, dans cette section, surtout de l'hypothèse de Poincaré en dimension > 3 ; on reviendra sur le cas 3-dimensionnel, dans les sections III et IV ci-dessous. Henri Poincaré énonce sa conjecture, "toute variété 3-dimensionnelle fermée, simplement connexe Σ^3 est la sphère 3-dimensionnelle S^3" dans son V$^{\text{ème}}$ Complément à la série de mémoires sur l'Analysis Situs [49], l'acte fondateur de ce qui est devenu, depuis, "la topologie". En fait, quelques années plus tôt, Poincaré croyait encore qu'une M^3 fermée avec le premier nombre de Betti nul, était S^3. Mais il découvre l'existence des sphères d'homologie (autres que S^3), introduit à cette occasion le groupe fondamental et énonce, finalement, la "Conjecture de Poincaré " sous sa forme correcte.

Pour des raisons qui ne me sont pas très claires, cette question ne semble pas du tout intéresser les successeurs immédiats de Poincaré et c'est W. Hurewicz (et aussi J.H.C. Whitehead) qui reprennent la question dans les années trente. Hurewicz a des notions très claires sur les questions d'homotopie et formule "l'hypothèse de Poincaré généralisée", "toute sphère d'homotopie Σ^n est la sphère S^n." A partir de ce moment, l'hypothèse de Poincaré devient un ingrédient de première importance dans le développement historique de la topologie.

Un rôle fondamental dans le développement futur du sujet, sera joué par la théorie du "type d'homotopie simple", découverte par J.H.C. Whitehead aux environs de 1940 [69]. Whitehead introduit des techniques géométriques nouvelles, les dilatations, les collapsings, les voisinages réguliers, e.a.d.s. et aussi une technologie algébrique nouvelle, dont la K-théorie algébrique est finalement sortie. (Le groupe de Whitehead, où habite la torsion de Whitehead, est essentiellement notre K_1.) Avec ses méthodes, il peut prouver, par exemple, que si Δ^3 est une boule d'homotopie, alors pour p assez grand, $\Delta^3 \times B^p = B^{p+3}$. Je dois ouvrir ici une petite parenthèse. Dans cet article, je vais appeler *boule d'homotopie* n-dimensionnelle une variété contractile Δ^n telle que $\pi_1 \partial \Delta^n = 0$. Il se trouve que si $n = 3$ cette dernière condition est automatiquement satisfaite, et toute Δ^3 est de la forme $\Sigma^3 - \text{int}\, B^3$, où $B^3 \subset \Sigma^3$ est une vraie boule (différentiellement) plongée dans Σ^3. Donc prouver que $\Sigma^3 = S^3$ ou que $\Delta^3 = B^3$ sont des choses équivalentes. Manifestement aussi, si on s'intéresse aux S^n et Σ^n on doit s'intéresser aussi aux B^n et Δ^n ainsi qu'aux espaces euclidiens R^n et à leurs avatars homotopiques.

Si $n \geq 4$, il existe des variétés contractiles M^n avec $\pi_1 \partial M^n \neq 0$ (voir, par exemple [42] ou [30], où de telles variétés apparaissent comme facteurs non-triviaux de B^{n+1}) ; donc la restriction $\pi_1 \partial \Delta^n = 0$ est nécessaire, si $n \geq 4$. Par ailleurs les questions $\Sigma^n = S^n$ et $\Delta^n = B^n$ sont, essentiellement, équivalentes. J'ai mis ici le "essentiellement" parce qu'encore faut-il préciser le cadre où l'on se place (voir le (2.1) ci-dessous).

Arrivé à ce point de notre histoire on doit introduire un nouveau thème. Quand, disons, Hurewicz écrivait l'égalité hypothétique $S^n = \Sigma^n$ on peut raisonnablement penser qu'il ne se posait pas trop de problèmes sur ce que l'égalité en question pouvait signifier. Mais, dans les années quarante, il devient clair que, pour les variétés, il y a (au moins) trois contextes possibles, dans lesquels on peut travailler, en ordre de généralité, donc de pauvreté, croissante

il s'agit des catégories

(2.1) différentiable (DIFF), linéaire-par-morceaux (PL) et topologique (TOP)

En 1956 John Milnor [32] fait la découverte fondamentale que sur S^7 on peut mettre *plusieurs* (en fait exactement 28) structures DIFF toutes différentes entre elles ; ceci veut dire qu'au moins pour les dimensions où de telles "sphères exotiques" existent, on ne peut pas avoir l'égalité $\Sigma^n \underset{\text{DIFF}}{=} S^n$, en toute généralité. Le travail de Milnor, suivi bientôt par un travail de Michel Kervaire [25] qui montre l'existence de variétés TOP sans aucune structure DIFF, ouvre une nouvelle étape dans l'histoire de la topologie. Faute de place on ne pourra pas la suivre dans cet article, et on n'y reviendra qu'épisodiquement, quand on en aura besoin. Disons, aussi, que dans cet article on se cantonnera, surtout dans le cadre DIFF.

Maintenant, il y a un cas particulièrement important de l'hypothèse de Poincaré en dimension n quelconque, le "problème de Schoenflies". Il s'agit de la question suivante. On considère un plongement DIFF $S^{n-1} \overset{j}{\hookrightarrow} S^n$, qui sépare S^n en deux boules d'homotopie X^n, Y^n. La question est de savoir si elles sont, effectivement B^n. La réponse positive pour $n \leq 3$ (Alexander [1]) était connue déjà depuis le début des années 1920, mais tous les efforts ultérieurs pour étendre ceci à $n = 4$, $n = 5$, e.a.d.s. se sont révélés vains. En 1958, Barry Mazur [29], qui avait moins de vingt ans à l'époque, prouve par un procédé vraiment miraculeux, que

$$X^n - \{\text{un point } p \text{ de } \partial X^n\} \underset{\text{DIFF}}{=} B^n - \{\text{un point } q \text{ de } \partial B^n\} \tag{2.2}$$

ce qui, en pensant aux compactifiés d'Alexandrov donne, tout de suite

$$X^n \underset{\text{TOP}}{=} B^n\,. \tag{2.3}$$

(Remarquez, ici, l'hypothèse DIFF et la conclusion (2.3) qui n'est que TOP.) L'impact mathématique, mais aussi psychologique de ce résultat est énorme. Toutes les dimensions y sont traitées, à la fois, sans que la difficulté augmente, justement, quand la dimension s'accroit.

Le pas suivant est franchi par Steve Smale qui a l'idée audacieuse que les grandes dimensions sont *plus faciles*. En 1960 il prouve [57] que

$$\Sigma^n \underset{\text{PL}}{=} S^n \quad \text{si} \quad n \geq 5\,. \tag{2.4}$$

En fait (2.4) suit d'un théorème plus général qui dit que tout h-cobordisme $(W^{n+1}; M_0^n, M_2^n)$, simplement connexe, avec $n+1 \geq 6$ est, $W^{n+1} \underset{\text{DIFF}}{=} M_0^n \times [0,1] \underset{\text{DIFF}}{=} M_1^n \times [0,1]$. On dit qu'un tel triple W^{n+1}, avec $\partial W^{n+1} = M_0^n + M_1^n$ est un h-cobordisme, si les inclusions $M_0^n \hookrightarrow W^{n+1} \hookleftarrow M_2^n$ sont des équivalences d'homotopie. Ce théorème de h-cobordisme simplement connexe de Smale [58] (et le cadre du présent article ne nous permettra pas de parler des aspects non-simplement connexes de la question, où la torsion de Whitehead intervient) implique que

$$\Delta^n \underset{\text{DIFF}}{=} B^n \quad \text{si} \quad n \geq 6\,. \tag{2.5}$$

Si pour Δ^5 on impose, aussi que $\partial\Delta^5 = S^4$ et si, de plus, on invoque le travail de Kervaire-Milnor [26], on peut étendre (2.5) pour tous les $n \geq 5$. Ceci implique (2.4), c'est-à-dire l'hypothèse de Poincaré en dimension $n \geq 5$. Du coup, aussi, on peut améliorer le (2.3) de Schoenflies, en

$$X^n \underset{\text{DIFF}}{=} B^n \text{ , au moins si } n \neq 4\,. \tag{2.6}$$

Le travail de Smale introduit une technologie nouvelle, les anses (handles) inspirée de la théorie de Morse, en topologie géométrique.

A peine Steve Smale avait-il annoncé ses résultats, John Stallings propose une méthode complètement différente [59], basée sur un très subtil "théorème d'engouffrement" (engulfing) qui lui permet de re-prouver $\Sigma^n \underset{\text{TOP}}{=} S^n$ pour $n \geq 5$ (légèrement plus faible que 2.4) mais, aussi un résultat concernant les variétés ouvertes [60] : Si V^n, avec $n \geq 5$ est ouverte, contractile et simplement connexe à l'infini, alors

$$V^n \underset{\text{DIFF}R}{\overset{n}{=}}\,. \tag{2.7}$$

On reviendra sur cette "simple connexité à l'infini", plus loin dans cet article, mais signalons que (2.7), combiné avec des résultats plus anciens pour $n \leq 3$ implique que [60]

(2.8) Si $n \neq 4$, R^n n'admet qu'une unique structure différentielle.

Pour être plus précis, le travail initial de Stallings ne prouvait $\Sigma^n \underset{\text{TOP}}{=} S^n$ que pour $n \geq 7$, les cas $n = 5$ et 6 étant prouvés tout de suite après par Christopher Zeeman [72]. Par ailleurs, le travail de Penrose-Whitehead-Zeeman [34] est à l'origine de la théorie de l'engulfing de Stallings.

Dans les paragraphes qui précèdent, je viens de passer très brièvement en revue la grande période héroïque de la topologie en grandes dimensions ; je ne suis pas seulement allé très vite sur l'aspect purement mathématique mais aussi sur l'aspect humain qui a accompagné cette histoire. Quoique aujourd'hui j'en connaisse très bien tous les protagonistes, à l'époque j'étais très loin, comme sur une planète étrangère ; je ne connais donc cet autre aspect que par ce qu'on m'a raconté plus tard. Il y aurait là tout un vrai roman à écrire, la réception d'un absolu et glacé scepticisme que le résultat de Barry Mazur a eu, les tâtonnements de Christopher Zeeman, suite aux très originales suggestions de son étudiant à l'époque, Roger Penrose, l'épisode Arbeitstagung (Bonn 1960) où pendant un moment l'approche de Steve Smale paraissait s'effondrer (exactement au point névralgique où elle touchait, via les anses d'indice $\lambda = 1$ aux groupes non-commutatifs) et seule l'approche de John Stallings restait en lice, e.a.d.s. Je pense qu'un acteur présent aux événements en question pourrait raconter tout ça mieux que moi.

Avant d'aller plus loin je reviens aux trois catégories (2.1) et à leurs inter-relations. Ceci n'est pas une parenthèse qui nous éloignerait de notre sujet principal ; je vais quand même être très bref, par manque d'espace.

Au début des années 1950 [54] il est déjà connu que, pour les dimensions $n \leq 3$, les trois catégories (2.1) sont essentiellement équivalentes. Mais l'un des plus spectaculaires succès de la topologie en grandes dimensions est d'avoir complètement élucidé en termes "d'espaces

classifiants" et de "classes caractéristiques", les relations entre les catégories TOP et PL en dimension $n \geq 5$. La relation entre PL et DIFF est relativement plus facile et la théorie qui en rend compte est une théorie d'obstructions à valeurs dans les groupes (abéliens et finis) de (classes de h-cobordisme des) sphères d'homotopie lisses de Kervaire-Milnor [26].

L'idée, très surprenante à l'époque, que la fameuse question du Hauptvermutung (unicité de la structure PL sur une variété TOP) soit accessible à une théorie d'obstruction, donc aux méthodes de la topologie algébrique, est due à Dennis Sullivan [62]. Son travail est suivi par un autre, tout aussi spectaculaire, de Rob Kirby et Larry Siebenmann [28]. Ce que Sullivan et Kirby-Siebenmann nous apprennent, par exemple, est qu'en grandes dimensions ($n \geq 5$) le Hauptvermutung, même si il n'est pas toujours vrai, est complètement contrôlé par les invariants discrets de la topologie algébrique. On aura à revenir sur ces questions dans la section IV ci-dessous.

Pour l'hypothèse de Poincaré , la prochaine étape importante est le travail de Michael Freedman [17] qui, en 1982 prouve que

$$\Sigma^4 \underset{\text{TOP}}{=} S^4 \,. \tag{2.9}$$

Cette percée très spectaculaire, la preuve de la Conjecture de Poincaré TOP, en dimension quatre, pour laquelle des résultats antérieurs d'Andrew Casson [6] sont essentiels s'accompagne, toujours dans la catégorie TOP, du théorème de h-cobordisme ($\pi_1 = 0$) en dimension $n = 5$ et d'une classification complète des variétés TOP, fermées, simplement connexes de dimension $n = 4$. Tout en utilisant une technologie beaucoup plus difficile que celles des grandes dimensions, Freedman nous dit qu'en dimension $n = 4$, dans la catégorie TOP, la relation entre géométrie et topologie algébrique est tout aussi directe qu'en grandes dimensions. Pour $\pi_1 = 0$, la forme quadratique d'intersection décide tout et toutes les formes unimodulaires sont réalisables, à ceci près que si la forme en question est impaire il lui correspond exactement deux variétés séparées par l'invariant $\alpha(M^4) \in Z/2Z$ de Kirby-Siebenmann ($\alpha(M^4) = 0 \Leftrightarrow M^4 \times S^1$ admet une structure DIFF). Les démonstrations sont un formidable enchevêtrement de processus infinis, d'une complexité beaucoup (infiniment ?) plus grande que ce que le sujet avait connu avant.

Ceci dit Freedman, utilise, entre autres, des techniques de la topologie "sauvage" (wild topology) à la Bing (et Bob Edwards), en particulier des arguments de "shrinking" qui n'ont aucune chance de passer à la catégorie DIFF (ou PL). Son travail est très solidement cantonné dans le TOP.

Mais presque tout de suite après Freedman, la topologie DIFF 4-dimensionnelle a elle aussi été bouleversée par Simon Donaldson [11] dont le travail met des restrictions extrêmement sévères sur ce que la forme quadratique d'intersection d'une variété fermée 4-dimensionnelle *différentiable* peut être. Tandis que Freedman montre qu'il y a beaucoup (disons "autant qu'on peut raisonnablement en vouloir") de variétés 4-dimensionnelles TOP, Donaldson montre que seulement une partie infime de l'ensemble de ces variétés possèdent des structures DIFF mais aussi que, généralement parlant, celles qui en possèdent en ont un très grand nombre. Ici les choses ne sont plus du tout comme en grandes dimensions mais plutôt comme en théorie des surfaces algébriques complexes. En catégorie DIFF, le théorème de h-cobordisme de Freedman devient faux et, résultat très spectaculaire obtenu en mettant ensemble les résultat de Donaldson et de Freedman [18], il existe des R^4 exotiques, résultat

qu'on contrastera avec (2.8). En fait, Clifford Tanks prouve qu'il y en a une infinité non-dénombrable [65].

Du coup, le Hauptvermutung n'est pas seulement faux en dimension $n = 4$, il est même si outrageusement faux que ses violations ne sont plus contrôlées par des invariants discrets ; la topologie algébrique est complètement incapable de rendre compte de ce qui se passe en dimension $n = 4$, dans la catégorie DIFF. En fait, le travail de Donaldson procède d'une partie complètement inattendue de notre science : les équations de Yang-Mills de la théorie de jauge, généralisation non-linéaire des équations de Maxwell . Du point de vue géométrique, la linéarité des équations de Maxwell (respectivement la non-linéarité des équations de Yang-Mills) tient au fait qu'elles sont liées à un fibré en droites complexes dont le groupe structural $U(1)$ est abélien (respectivement au fait qu'elles sont liées à un fibré en plans complexes dont le groupe structural $SU(2)$ est non-abélien). Du point de vue physique cette linéarité tient au fait que les photons, les quanta du champ électromagnétique n'interagissent pas entre eux, respectivement au fait que les bosons intermédiaires de la théorie de Yang-Mills interagissent entre eux.

La grande découverte de Donaldson est que des renseignements extrêmement précieux concernant la topologie différentielle de M^4, peuvent être obtenus via l'étude de l'espace des modules des solutions de Yang-Mills sur M^4.

Il n'y a pas lieu de décrire ici, ni les remarquables nouveaux invariants découverts par Donaldson, ni la très surprenante "révolution de novembre" (1994) quand Nathan Seiberg et Edward Witten proposent un système de nouvelles équations couplées, d'une remarquable efficacité, et dont on n'a pas fini encore de tirer toutes les conséquences.

Ce qu'il faut retenir de tout ceci, en premier lieu, pour notre sujet présent est qu'exactement en dimension $n = 4$, entre les mondes DIFF et TOP il y a un énorme précipice dont les techniques topologiques classiques ne peuvent aucunement rendre compte et que, par ailleurs, on est encore très loin de bien comprendre.

Même avec toutes les percées spectaculaires en topologie différentielle de dimension quatre (Donaldson, Taubes, Seiberg-Witten, Kronheimer-Mrowka,...) les questions fondamentales suivantes restent encore complètement ouvertes (et on peut raisonnablement penser, demanderont encore une technologie nouvelle).

A) Est-ce que $\sum^4 \underset{\text{DIFF}}{=} S^4$? Une forme équivalente de la question est la suivante : si Δ^4 est une boule d'homotopie avec $\partial\Delta^4 = S^3$, est-il vrai que $\Delta^4 = B^4$? Ici, pour prouver l'équivalence en question, on a besoin d'invoquer le théorème "$\Gamma_4 = 0$" de Jean Cerf [7].

B) En particulier, existe-t-il des B^4 exotiques ? Une boule de Schoenflies différentiable de dimension $n = 4$ est-elle toujours standard ? (La réponse "oui" à cette dernière question voudrait dire que (2.6) est vrai sans restriction dimensionnelle.)

C) Pour une boule d'homotopie 4-dimensionnelle tout à fait générale, a-t-on toujours $\Delta^4 \underset{\text{DIFF}}{=} B^4$? (La réponse "oui" entrainerait, aussi, l'hypothèse de Poincaré en dimension 3, car $\sum^3 \# (-\sum^3)$ est h-cobordant à S^3.)

Avec tout ceci, le moment est venu, justement, pour passer à la dimension $n = 3$, qui va nous occuper pour les trois prochaines sections de cet article.

3 Un très bref aperçu sur le programme de Thurston

L'un des plus importants résultats d'Henri Poincaré est son fameux théorème d'uniformisation pour les surfaces de Riemann ($n = 2$). Ce théorème, qui l'occupe entre 1882 et 1907 est, comme on le sait, l'objet d'une compétition assez acharnée avec Klein et Koebe.

Dans les années soixante dix, William Thurston a formulé une vaste conjecture de géométrisation, disons l'analogue du théorème d'uniformisation de Poincaré , en dimension $n = 3$ [66]. La conjecture de géométrisation implique l'hypothèse de Poincaré , et beaucoup plus. Le cadre de cet article ne me permet pas de décrire les huit géométries 3-dimensionnelles découvertes par Thurston ni donc de formuler la conjecture de géométrisation dans sa version complète. Je vais donc me limiter ici à une version simplifiée, que voici.

Conjecture de Géométrisation de Thurston (ou "Uniformisation géométrique", en version simplifiée)

"Une variété fermée irréductible M^3 vérifie l'une des trois conditions, mutuellement exclusives, suivantes

i) M^3 admet une métrique riemannienne avec courbure sectionnelle $K = -1$.

ii) M^3 admet une métrique riemannienne avec courbure sectionnelle $K = +1$.

iii) $Z + Z \subset \pi_1 M^3$."

Des gros morceaux de cette conjecture ont déjà été démontrés par Thurston [66] et d'autres mathématiciens, avec des conséquences spectaculaires comme la résolution de la conjecture de Smith [31].

4 Le programme de Po pour l'hypothèse de Poincaré

Je vais donner, dans la présente section IV une succinte présentation de mon propre programme pour l'hypothèse de Poincaré, le "programme de Po" [13]. Les premiers paragraphes de la section en décrivent les motivations générales.

Toute sphère d'homotopie Σ^3 peut être reliée à S^3 par un chemin "continu" d'objets *singuliers* de dimension $n = 3$. Whitehead avait déjà proposé de tels chemins et il savait, aussi, les rendre lisses en grandes dimensions. Mais c'est seulement si le chemin peut être rendu lisse en dimension $n = 4$ et pas plus, qu'on peut en tirer la conclusion désirée que $\Sigma^3 = S^3$.

Ceci demande, pour commencer, un chemin singulier d'une très grande sophistication. Les [37], [38] développent un calcul (the honeycomb calculus) qui engendre de tels chemins. Bien entendu, il y a des obstructions combinatoires très précises à la lissification 4-dimensionnelle. En fait, ces obstructions (obstructions to "COHERENCE") ont des propriétés de conservation qui font qu'aucune démarche de type fini, même avec des sophistications arbitrairement grandes du calculus, ne peuvent jamais les annuler. Une fois qu'on a réalisé qu'on doit vivre avec les obstructions, l'idée générale est d'*itérer infiniment* les opérations du calculus. A chaque étape finie les obstructions restent présentes, mais l'idée est de conduire ce *processus infini* de telle façon que, dans un sens convenable, il converge. *A la limite*, on obtient un objet lisse qui sera une variété différentielle 4-dimensionnelle *non-compacte* et à bord non-vide, à partir de laquelle on puisse tirer des renseignements intéressants pour notre problème.

En 1966 John Stalling [61], en parlant des difficultés de l'hypothèse de Poincaré disait "... we always run against the problem of trying to simplify, by some geometric trick the situation. But any little homotopy that would simplify the picture always, in fact, greatly complicates it." De ce point de vue, l'idée du programme est de reconnaître, pour commencer que les difficultés ou obstructions ne peuvent pas être annulées et, à partir de là, de les repousser à l'infini.

Après cette brève introduction générale au programme je vais passer à un niveau de description plus précis.

On va dire qu'une variété M^n est *géométriquement simplement connexe*, en abréviation on écrira souvent g.s.c., si elle possède une décomposition en anses (= handlebody decomposition), PROPRE, sans anses d'indice $\lambda = 1$. Si $\partial M^n = \emptyset$ cette condition s'exprime, aussi, en disant qu'il y a une fonction de Morse PROPRE $M^n \xrightarrow{f} R_+$, sans singularités d'indice $\lambda = 1$; pour le cas où $\partial M^n \neq \emptyset$ on peut encore donner une formulation fonctionnelle, en insistant, pour commencer, que toutes les singularités de f soient dans $\operatorname{int} M^n$ et en incluant ensuite aussi celles des singularités de $f \mid \partial M^n$ qui sont "essentielles" (dans le sens que le passage $f^{-1}(-\infty, c-\varepsilon] \Rightarrow f^{-1}(-\infty, c+\varepsilon]$ pour la valeur critique respective c, entraîne un changement de l'homologie ; pour les singularités "non-essentielles", ce passage est sans saut homologique). Ceci dit, il est évident que M^n g.s.c. entraîne que $\pi_1 M^n = 0$, mais en ce qui concerne l'implication inverse, la situation est beaucoup plus complexe ; elle dépend de deux paramètres indépendants, de la dimension n et aussi de la compacité ou non-compacité de M^n. Voici un très bref aperçu de ce qui se passe, dans le cas compact, pour commencer. Si $n \geq 5$, avec $\partial M^n = \emptyset$ ou $\partial M \neq \emptyset$, on a l'implication $\pi_1 M^n = 0 \Rightarrow M^n$ g.s.c. ; ceci résulte des techniques introduites par Smale [58] pour le théorème de h-cobordisme. Si $n = 4$, une obstruction subtile découverte par A. Casson [14], liée aux représentations du $\pi_1 \partial M^4$, peut interdire que notre implication soit vérifiée. Par exemple, l'obstruction de Casson est non-nulle pour les facteurs non-triviaux de B^5, découverts par Po [42] et par Barry Mazur [30] à la fin des années cinquante ; ces variétés contractiles ne sont pas g.s.c. Enfin, pour $n = 3$, et on reste ici encore dans le contexte compact, l'implication $\pi_1 M^3 = 0 \Rightarrow M^3$ g.s.c. est exactement équivalente à l'hypothèse de Poincaré , le sujet de cette conférence. Si M^n est ouverte (c'est-à-dire non-compacte, avec $\partial M^n = 0$), et si $n \geq 5$, alors il n'est pas très difficile de voir que $\pi_1 M^n = 0$ entraîne M^n g.s.c., à condition que M^n soit, aussi, simplement connexe à l'infini (notion qu'on va complètement expliciter un peu plus loin) ; une preuve de ce résultat se trouve, par exemple, dans [50]. En revanche, quand M est non-compacte, à bord $\partial M \neq \emptyset$, et il se trouve que c'est exactement la situation qui va beaucoup nous intéresser dans ce qui suit, alors il n'y a pas, à ma connaissance, des résultats généraux d'énoncés aussi simple que ceux qu'on vient de passer en revue.

Après avoir éclairci ainsi la notion de simple connexité géométrique, voici un autre petit préliminaire : en dimension $n = 3$, à toute sphère d'homotopie Σ^3 s'associe canoniquement la boule d'homotopie $\Delta^3 = \Sigma^3 - \operatorname{int} B^3$, obtenue à partir d'un plongement lisse $B^3 \subset \Sigma^3$ et, réciproquement, à toute boule d'homotopie Δ^3 on associe canoniquement une sphère d'homotopie Σ^3, puisque $\partial \Delta^3 = S^2$, et on peut recoller un B^3 le long de ce $\partial \Delta^3$. Ainsi, l'hypothèse de Poincaré peut être reformulée "$\Delta^3 = B^3$" et, dans la suite, on y pensera sous cette forme.

Voici, maintenant, les trois étapes principales du programme de Po, tel qu'il se présente

à l'heure où j'écris ces lignes. (D'autres présentations du programme se trouvent, aussi, dans [13], que je recommande vivement, ou [43], [44], mais ce qui suit est plus à jour.)

Étape I. Le résultat principal est ici le suivant

Théorème A. *Pour toute boule d'homotopie de dimension trois Δ^3, la variété lisse ouverte de dimension $n = 4$,*

$$X^4 = \text{int}\,((\Delta^3 \times I) \# (S^2 \times D^2) \# (S^2 \times D^2) \# \ldots) \tag{4.1}$$

est g.s.c.

Ici "#" veut dire somme connexe le long du bord et, sauf mention explicite du contraire, c'est ainsi que cette notation sera entendue dans la suite. Le "$\# (S^2 \times D^2) \# (S^2 \times D^2) \# \ldots$", qu'on notera aussi par $\# \infty \# (S^2 \times D^2)$, est une notation qui comporte une certaine ambiguité, vu qu'elle peut recouvrir une infinité (non dénombrable) d'espaces des bouts (ends) tous différents les uns des autres. Mais le "int" par lequel commence la définition (4.1) de X^4, et qui efface la frontière, fait disparaître, aussi, toute ambiguité sur les bouts. La démonstration complète du Théorème A est couverte par les articles suivants [36], [37], [38], [[**?**], parties A et B], [40].

Remarque. Je vais expliquer maintenant comment des objets 4-dimensionnels du type (4.1) apparaissent dans le contexte qui nous intéresse ici.

Considérons donc, pour une boule d'homotopie Δ^3, et pour un entier positif p, la variété 4-dimensionnelle lisse, compacte, à bord non-vide

$$X_0^4 = (\Delta^3 \times I) \# p \# (S^2 \times D^2) \tag{4.2}$$

qui ressemble au X^4 de (4.1), mais en beaucoup plus simple. Je dis que *si* l'on a un difféomorphisme

$$X_0^4 = B^4 \# p \# (S^2 \times D^2)\,, \tag{4.3}$$

alors $\Delta^3 = B^3$, la chose qu'on désire. En effet, en prenant les bords des deux termes de (4.3) on trouve

$$2\,(\Delta^3 \# p \# (S^2 \times I)) = 2\,(B^3 \# p \# (S^2 \times I))\,; \tag{4.4}$$

ici, si V^3 est une variété à bord $\partial V^3 \neq \emptyset$, on note par ∂V^3 la variété (sans bord) obtenue en recollant deux exemplaires de V^3 le long des ∂V^3. Le membre droit de (4.4) est un objet "standard", dont le revêtement universel se plonge dans R^3. On en tire un plongement lisse $\Delta^3 \subset R^3$ qui, par Alexander [1] implique que $\Delta^3 = B^3$. Le théorème d'Alexander dont il est question ici, est le cas $n = 3$ du problème de Schoenflies, déjà mentionné.

Le lecteur remarquera que si l'on remplace l'assertion (à ce jour purement conjecturale) (4.3) par un énoncé similaire en une dimension $n > 4$ (et ces énoncés sont connus pour être vrais, voir par exemple [58]) la démarche qu'on vient d'esquisser, c'est-à-dire l'implication (facile)

$$(4.3) \Rightarrow \Delta^3 = B^3 \text{ (Hypothèse de Poincaré)}$$

ne fonctionnerait plus du tout. D'où l'intérêt, pour le programme, une fois qu'on a admis qu'on ira en dimension $n > 3$, à se cantonner en dimension $n = 4$, et pas plus.

Étape conjecturale II. Il s'agit ici de prouver le résultat suivant

Conjecture B. *Si X^4 est comme dans* (4.1)*, alors on a l'implication suivante*

$$X^4 \text{ est g.s.c. } \Longrightarrow \Delta^3 \times I \text{ est g.s.c.} \tag{4.5}$$

(Les prémices de (4.5) est la conclusion du Théorème A, mais ce qu'il s'agit de prouver ici est l'*implication*, c'est-à-dire la simple connexité géométrique de $\Delta^3 \times I$.) Cette étape conjecturale fait l'objet d'un long travail de collaboration avec David Gabai. On a commencé cette collaboration en 1995, quand une faille a été découverte dans une première version du programme (je vais revenir là-dessus plus loin). Les techniques qu'on a développé depuis, semblent (à ce jour) pouvoir s'appliquer à d'autres situations, aussi, en particulier au problème de Schoenflies DIFF en dimension $n = 4$. Mais, encore une fois, il s'agit ici de recherches en cours et on ne va pas beaucoup en parler ici.

Étape III. Maintenant, le résultat principal est le

Théorème C. *Soit Δ^3 une boule d'homotopie 3-dimensionnelle, telle que $\Delta^3 \times I$ soit g.s.c. Alors, on a $\Delta^3 = B^3$.*

En d'autres termes, ce théorème nous dit que ce qu'on appelle la "Conjecture de Poincaré stable", implique la conjecture de Poincaré tout court. Par ailleurs, il est clair que la juxtaposition des théorèmes A et C, et aussi de l'énoncé conjectural B, impliquent que $\Delta^3 = B^3$ pour tout Δ^3, c'est-à-dire impliquent la Conjecture de Poincaré . La preuve du Théorème C est complètement contenue dans [41].

On va regarder maintenant le programme d'un peu plus près. Il faut dire, tout d'abord, que les démarches qui prouvent les deux théorèmes A et C sont apparentées entre elles. Les deux utilisent des processus infinis, conduisant à des variétés lisses, de dimension $n = 4$, non-compactes et à bord non-vide.

Typiquement ces variétés qu'on notera génériquement par V^4, rentrent dans le cadre de la description générale que voici

$$\operatorname{int} V^4 \underset{\text{DIFF}}{=} R^4_{\text{standard}}\,,\ \partial V^4 = \sum_1^\alpha S^1_i \times \operatorname{int} D^2_i\,. \tag{4.6}$$

Ici on insiste sur le fait que différentiablement parlant, $\operatorname{int} V^4$ est un R^4 qui n'est pas exotique (rappelez-vous qu'exactement en dimension $n = 4$, via Freedman et Donaldson, cette "pathologie" peut arriver). Par ailleurs, $1 \leq \alpha \leq \infty$ et la trivialisation (= framing) $\sum\limits_i S^1_i \times \operatorname{int} D^2_i$ du bord est "nulle", dans le sens suivant. Si on utilise la trivialisation en question pour changer chaque section nulle $S^1_i \times 0$ en $S^1_i \times \varepsilon$, et si l'on considère deux 2-chaines orientées C^2_0, C^2_ε de V^4 bordant ces deux cercles, alors le nombre d'intersection algébrique $C^2_0 \cdot C^2_\varepsilon = 0$.

Je vais appeler les variétés V^4 du type qu'on vient de décrire, des "*espèces de link*" (sort of links) ou, des "*espèces de nœuds*" (sort of knots) si $\alpha = 1$ dans (4.6). Cette terminologie est motivée par la recette suivante qui permet de construire (certains) des V^4. Prenons

$1 \leq \alpha < \infty$ et partons d'un link classique L, c'est-à-dire d'un plongement lisse

$$\sum_{1}^{\alpha} S_i^1 \underset{L}{\hookrightarrow} S^3 \,. \tag{4.7}$$

Pour L on considère la trivialisation nulle (ou de Seifert, dans le jargon de la topologie 3-dimensionnelle)

$$X^3 \underset{\text{def}}{=} \sum_{1}^{\alpha} S_i^1 \times \operatorname{int} D_i^2 \hookrightarrow S^3 = \partial B^4_{\text{standard}} \,. \tag{4.8}$$

[1, En toute dimension $n \neq 4$, il est connu que toute variété lisse homéomorphe à B^n est différentiablement standard ; pour $n > 4$ ceci résulte du théorème de h-cobordisme de Smale, tandis que pour $n \leq 3$ il y a équivalence de catégories DIFF $\approx$ TOP. Mais, si $n = 4$, il n'est pas connu à ce jour si des B^4 lisses *exotiques* existent. Dans la présente discussion on insiste que les B^4 lisses dont on parle soient, justement, standard.]vec tout ceci, on passe à la variété suivante qui, manifestement, est une espèce de link

$$V^4 \underset{\text{def}}{=} B^4 - (\partial B^4 - X^3) \,, \tag{4.9}$$

où $B^4 = B^4_{\text{standard}}$ est la boule 4-dimensionnelle cobord du S^3 de (4.7), (4.8). On dira qu'une espèce de link V^4 qui, à difféomorphisme près, se laisse décrire par ce procédé, c'est-à-dire par des formules (4.7) + (4.8) + (4.9), est "*différentiablement non-sauvage*" (smoothly tame). En d'autres termes encore, V^4 est différentiablement non-sauvage si elle admet une compactification lisse en un B^4_{standard}. Sinon, on dira que V^4 est différentiablement *sauvage* (smoothly wild). Mais il y a, aussi, une notion plus large que la non-sauvagerie différentiable : on dira que V^4 est topologiquement non-sauvage (topologically tame) si elle admet une description du type (4.9) à homéomorphisme près, seulement. On touche ici, de plein pied à la distinction entre catégories TOP et DIFF, en dimension $n = 4$. Pour être un peu plus spécifique, dans l'approche de M. Freedman à la Conjecture de Poincaré TOP en dimension $n = 4$ [17] (voir, aussi, [19]) un rôle crucial est joué par les *anses de Casson* (= Casson Handles) [6] qui, typiquement, sont des espèces de nœuds, dans le sens qu'on a défini plus haut, mais d'un type très très particulier. Dans [17] le pas fondamental franchi par Freedman (par une énorme symphonie de processus infinis, Larry Siebenmann dixit) est la preuve du fait que toute anse de Casson (on écrira CH, en abrégé) est topologiquement standard (c'est-à-dire que CH $\underset{\text{TOP}}{=} D^2 \times \operatorname{int} D^2$), donc automatiquement topologiquement non-sauvage. Mais, en combinant les résultats de M. Freedman et ceux de S. Donaldson [11] (voir, aussi [18], [27]) on peut prouver qu'il existe des CH tels que CH $\underset{\text{DIFF}}{\neq} D^2 \times \operatorname{int} D^2$ et, il n'est pas du tout difficile de montrer que de tels CH sont, forcément, différentiablement sauvages. Disons, aussi, que si des CH non-sauvages n'existaient pas, il n'y aurait pas de R^4 exotiques non plus.

Tout ceci ayant été dit, l'un des ingrédients majeurs pour l'étape I du programme est le THÉORÈME DE NON-SAUVAGERIE LISSE (= smooth tameness theorem), que je vais présenter maintenant. Dans ce théorème apparaissent des espèces de link à une infinité de

composantes connexes du bord ($\alpha = \infty$, dans le contexte de (4.6))

$$\partial V^4 = \sum_1^\infty S_i^1 \times \operatorname{int} D_i^2 \,. \tag{4.10}$$

Pour un tel V^4, si $N \in Z_+$ est fixé, on définit la *troncation* finie (finite truncation)

$$V^4 \mid N = V^4 - \sum_{N+1}^\infty S_i^1 \times \operatorname{int} D_i^2 \,. \tag{4.11}$$

Notons, aussi, qu'en tout état de cause, le bord paramétré de V^4 est une invitation explicite à ajouter à V^4 des anses de dimension $n = 4$ et d'indice $\lambda = 2$, privées de leurs surfaces latérales, c'est-à-dire de passer du V^4 (qu'on considère maintenant avec $\alpha = \infty$) à la variété ouverte

$$V^4 \cup \sum_1^\infty D_i^2 \times \operatorname{int} D_i^2 \,, \tag{4.12}$$

où chaque $D_i^2 \times \operatorname{int} D_i^2$ est recollé à V^4 suivant la recette suivante

$$\partial V^4 \supset S_i^1 \times \operatorname{int} D_i^2 = \partial D_i^2 \times \operatorname{int} D_i^2 \subset D_i^2 \times \operatorname{int} D_i^2 \,.$$

Avec tout ceci, voici le

Théorème D. (Théorème de non-sauvagerie lisse.) *Pour toute boule d'homotopie Δ^3 on peut trouver une espèce de link V^4, avec les deux propriétés suivantes.*

1) *On a un difféomorphisme*

$$\operatorname{int}\left((\Delta^3 \times I) \# \infty \# (S^2 \times D^2)\right) = V^4 \cup \sum_1^\infty D_i^2 \times \operatorname{int} D_i^2 \,. \tag{4.13}$$

Ici, le premier terme est de X^4 de (4.1) *et le second est le même que* (4.12).

2) *Quel que soit $N < \infty$, la troncation $V^4 \mid N$ est différentiablement non-sauvage.*

Le V^4 du théorème ci-dessus est, a priori, pas du tout unique, pour Δ^3 fixé. On construit cette espèce de link par un *processus infini* très complexe, qui rassemble une infinité de pièces détachées 4-dimensionnelles compactes, chacune emboîtée dans la suivante, avec des π_1 de plus en plus gros, mais chacun de ces π_1 étant tué quand on monte suffisamment haut dans la construction, pour obtenir "à la limite" un objet non-compact et simplement connexe (une toute petite idée du genre de choses dont je parle, mais dans un cadre infiniment plus simple, est constitué par le petit "jouet" (toy model) (4.14), (4.15) ci-dessus).

Maintenant, une fois qu'un $N < \infty$ a aussi été fixé, la non-sauvagerie lisse de $V^4 \mid N$ est prouvée par un second processus infini, qui rassemble à l'infini de $V^4 \mid N$ des morceaux 3-dimensionnels, compacts, toujours non-simplement connexes, en nombre infini, jusqu'à ce qu'on ait étendu le bord $\partial(V^4 \mid N)$ en un bord plus grand, *connexe* et difféomorphe à R^3. A partir de cette extension de $V^4 \mid N$ on peut déduire la non-sauvagerie lisse du même $V^4 \mid N$ en invoquant le résultat relativement élémentaire que voici

Toute variété lisse W^4*, telle que* $\operatorname{int} W^4 = R^4_{\text{standard}}$ *et* $\partial W^3 = R^3$ *est, différentiablement parlant standard (c'est-à-dire que* $W^4 \underset{\text{DIFF}}{=} R^4_+$.

Il faut signaler ici, aussi, que l'analogue 3-dimensionnel du petit résultat qu'on vient d'énoncer se trouve, via Artin-Fox [2], être *faux*.

Avant de reprendre le fil de notre discussion principale et discuter le Théorème D je vais ouvrir une longue parenthèse qui, d'une part pourra aider à mieux comprendre le "processus infini" et "espèces de link" qu'on a évoqués ci-dessus mais qui introduira aussi les notions de *simple connexité à l'infini* et *groupe fondamental à l'infini* π_1^∞, dont il sera beaucoup question dans la suite.

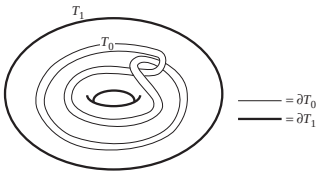


FIG. 4 – $T_0 \subset \operatorname{in} T_1$

Pour commencer, on propose comme "toy-model", un processus infini, très classique d'ailleurs, beaucoup moins sophistiqué que ceux du programme. La Figure 1 montre deux exemplaires du tore plein $S^1 \times D^2$, appelons-les T_0, T_1, liés par un plongement $T_0 \underset{i}{\subset} \operatorname{int} T_1$ qui est homotopiquement trivial mais pas isotopiquement trivial. On peut itérer indéfiniment le plongement i, c'est-à-dire plonger $T_1 \subset \operatorname{int} T_2$, exactement de la même manière qu'on a plongé T_0 dans T_1, e.a.d.s. Si, dans la suite infinie

$$T_0 \subset T_1 \subset T_2 \subset T_3 \subset \cdots \tag{4.14}$$

on s'arrête (à un cran fini) on a un T_n avec $\pi_1 T_n = Z$, tandis que la variété ouverte

$$\mathrm{Wh}^3 \underset{\text{def}}{=} \bigcup_{n=0}^{\infty} T_n \tag{4.15}$$

est contractile. Ce Wh^3, appelé *la variété de Whitehead*, est découvert par ce dernier dans les années trente [67]. Pour l'histoire, un an plus tôt Whitehead pensait avoir prouvé que toute variété ouverte 3-dimensionnelle contractile est R^3, ce qui aurait impliqué l'hypothèse de Poincaré. Mais, justement, Wh^3 est un contre-exemple à cette affirmation plus ambitieuse car, quoique contractile Wh^3 n'est *pas* homéomorphe à R^3. Comme Whitehead s'en rend compte, on peut trouver une suite de lacets λ_n de Wh^3 tels que $\lambda_n \to \infty$ dans Wh^3, sans que l'on puisse trouver des disques (singuliers) D^2_n de Wh^3 avec $\partial D^2_n = \lambda_n$ et $D^2_n \to \infty$.

On touche là à un thème très important pour toute mon approche (sections IV et V de cet article) et on va commencer par formaliser les affirmations plus ou moins heuristiques que je viens de faire, sous la forme suivante. On dira qu'un espace localement compact E

est *simplement connexe à l'infini*, si pour tout compact $k \subset E$ il existe un compact plus grand $k \subset K \subset E$ tel que tout lacet $\lambda \subset E - K$ qui est homotope à zéro dans E, l'est déjà dans $E - k$. Avec ceci, R^n est simplement connexe à l'infini (sauf si $n = 2$, manifestement), tandis que Wh^3 ne l'est pas (ce n'est pas dur à prouver).

Sous certaines conditions on peut définir, aussi, le *groupe fondamental à l'infini* de E, disons $\pi_1^\infty E$ qui, en principe, est la limite projective

$$\varprojlim_{K \subset E,\, \text{compact}} \pi_1(E - K)\,,$$

si et quand on peut lui donner un sens (il y a des problèmes de choix des points base, e.a.d.s.). En principe, toujours, la formule ci-dessus fait du $\pi_1^\infty E$ un *pro*-groupe, un groupe muni d'une topologie qui peut s'avérer très compliquée.

Ces (pro)-groupes π_1^∞ vont intervenir dans l'étape III du programme, dont on parlera plus loin dans cette section.

Mais la variété de Whitehead est exemplaire, à plus d'un titre, pour tout le cercle d'idées autour desquelles tourne tout cet article. On va donc en dire un peu plus. On a remarqué que $\mathrm{Wh}^3 \neq R^3$ mais, en revanche, il est facile de montrer que

$$\mathrm{Wh}^3 \times (0,1) \underset{\mathrm{DIFF}}{=} R^4_{\text{standard}}\,. \tag{4.16}$$

En partant de cette remarque simple, c'est-à-dire du fait que

$$(\operatorname{int} T_0 \times (0,1)) \cup (\operatorname{int} T_1 \times (0,1)) \cup (\operatorname{int} T_2 \times (0,1)) \cup \cdots \underset{\mathrm{DIFF}}{=} R^4_{\text{standard}} \tag{4.17}$$

on peut *ajouter un bord* $S^1 \times \operatorname{int} D^2$ à $\mathrm{Wh}^3 \times (0,1)$, en remplaçant le premier terme $T_0 \times (0,1)$ de (4.17) par $T_0 \times [0,1)$. Le

$$V^4 \underset{\text{def}}{=} (\operatorname{int} T_0 \times [0,1)) \cup (\operatorname{int} T_1 \times (0,1)) \cup (\operatorname{int} T_2 \times (0,1)) \cup \cdots \tag{4.18}$$

est manifestement une espèce de nœud, dans le sens de (4.6) avec $\alpha{=}1$. On peut même prouver (voir [51]) que c'est une Anse de Casson (disons que dans ce que, heuristiquement, on peut appeler "l'espace des modules des recettes DIFF pour CH", espace de modules qui est un Cantor de R, c'est l'un des deux points extrêmaux). Ainsi, par Freedman on sait que le V^4 de (4.18) est, du point de vue TOP, le $D^2 \times \operatorname{int} D^2$ mais, par ailleurs, je soupçonne très fortement qu'il est différentiablement sauvage. Je vais revenir là-dessus plus loin, mais là je ferme la parenthèse.

En ce qui concerne le Théorème D, bien que son énoncé puisse sembler, a priori, plutôt ésothérique, ce théorème peut être utilisé très efficacement, ce qui va être illustré ci-dessous. Avant, on aura besoin d'introduire une nouvelle notion, plus faible que la simple connexité géométrique. Soit M^n une variété lisse quelconque ; les situations qui nous concernent sont, en fait, ou bien M^n ouverte, ou bien M^n compacte avec $\partial M^n \neq \emptyset$. On dira que M^n est *géométriquement simplement connexe à longue distance* (terminologie que je dois à Barry Mazur) si pour tout compact $K \subset \operatorname{int} M^n$, on peut trouver une sous-variété lisse compacte N^n qui soit g.s.c., telle que

$$K \subset N^n \subset M^n\,. \tag{4.19}$$

Cette définition a une certaine ressemblence formelle avec celle de "simplement connexe à l'infini" ($\pi_1^\infty = 0$) donnée plus haut. Il y a, en fait une certaine relation entre le s.c.g. et le $\pi_1^\infty = 0$, relation qu'on explicitera et qu'on exploitera, dans la section V de cet article.

On notera que pour M^n compacte, avec $\partial M^n \neq \emptyset$ il suffit de prendre un voisinage-collier du bord $\partial M^n \times [0,1] \subset M^n$, de l'utiliser pour retirer M^n dans son intérieur, disons de passer à un autre exemplaire $M^n_{\text{petit}} \subset \operatorname{int} M^n$ et de demander s'il existe un N^n comme ci-dessus, tel que $M^n_{\text{petit}} \subset N^n \subset M^n$.

Je peux énoncer maintenant le

Corollaire E. (du théorème de non-sauvagerie lisse) *Pour toute boule d'homotopie Δ^3, la variété 4-dimensionnelle $\Delta^3 \times I$ est g.s.c. à longue distance.*

Voici, en quelques mots comment on peut déduire ce résultat, à partir du Théorème D ; je pense que c'est un petit exercice instructif pour le lecteur, raison pour laquelle je le donne ici. Pour le X^4 de (4.1) on considère un compact $K \subset X^4$. D'après (4.13), on a $X^4 = V^4 \cup \sum_1^\infty (D_i^2 \times \operatorname{int} D_i^2)$ et, vu la compacité de K, il existe un $N < \infty$ tel que

$$K \subset (V^4 \mid N) \cup \sum_1^N D_i^2 \times \operatorname{int} D_i^2 \,. \tag{4.20}$$

Comme $V^4 \mid N$ est différentiablement non-sauvage le second membre de (4.20) est l'intérieur de la variété 4-dimensionnelle lisse, *compacte*, à bord non-vide

$$Z^4 = B^4_{\text{standard}} + \left\{ \text{les } N \text{ anses d'indice } \lambda = 2 \,,\ \sum_1^N D_i^2 \times D_i^2 \right\} . \tag{4.21}$$

Il est clair que ce Z^4 est g.s.c. et il n'est pas difficile, en retirant Z^4 très légèrement vers son intérieur, d'interpoler Z^4 entre K et X^4

$$K \subset Z^4 \subset X^4 \,. \tag{4.22}$$

Ceci prouve, pour le moment, que X^4 est g.s.c. à longue distance.

Soit maintenant

$$A^4_n \underset{\text{def}}{=} (\Delta^3 \times I) \# n \# (S^2 \times D^2) \,, \tag{4.23}$$

et soit $A^4_n \subset \operatorname{int} A^4_{n+1}$ le plongement standard. On a, manifestement, $X^4 = \bigcup_{n=0}^\infty A^4_n$. Maintenant, on peut prendre le compact $A^4_0 = \Delta^3 \times I \subset X^4$ et trouver un M^4 qui soit compact et g.s.c. tel que

$$A^4_0 \subset M^4 \subset X^4 \,.$$

Par compacité, il existe un $p < \infty$ tel que

$$A^4_0 \subset M^4 \subset A^4_p \,, \tag{4.24}$$

et comme l'inclusion $A^4_0 \subset A^4_p$ est standard, de (4.24) on tire

$$(\Delta^3 \times I)_{\text{petit}} \subset M^4 \subset (\Delta^3 \times I) \,, \tag{4.25}$$

ce qui implique que $(\Delta^3 \times I)$ est bien g.s.c. à longue distance.

Le Théorème A est la conséquence immédiate de la juxtaposition du Corollaire E avec le résultat beaucoup plus récent [40], que voici

Théorème F. *Soit* Δ^4 *une variété lisse* 4*-dimensionnelle, compacte à bord non-vide, telle que* $\partial\Delta^4$ *soit une sphère d'homologie. Si* Δ^4 *est g.s.c. à longue distance, alors la variété ouverte* $\operatorname{int}(\Delta^4 \# \infty \# (S^2 \times D^2))$ *est g.s.c.*

Avant de passer à la discussion de l'étape III du programme (Théorème C) je vais faire deux commentaires, de nature très diverse.

D'abord un petit point d'histoire (personnelle). Une faille a été découverte au printemps 1995 par David Gabai et Michael Freedman, dans une première version de mon programme. Cette faille concernait un morceau indépendant du reste, qu'on appellera "Po-V", et qui devait faire le passage direct du Corollaire E au Théorème C. Plus précisément, Po-V devait prouver, directement, l'implication $\{\Delta^3 \times I$ g.s.c. à longue distance (résultat du Corollaire E)$\} \Rightarrow \{\Delta^3 \times I$ g.s.c. (prémices du Théorème C)$\}$. Dans la présente version du programme le Po-V est remplacé par le Théorème F ("le Po-V-A") dont la preuve a été complétée à la fin de 1998, et par l'étape conjecturale B (que David Gabai et moi on appelle, entre nous, le "Po-V-B" ; on est en train de travailler dessus). Je n'en dirai pas plus sur l'étape II, maintenant.

Mais je veux expliquer pourquoi, dans un travail qui, finalement devrait concerner la dimension $n = 3$ où il y a équivalence entre les catégories DIFF et TOP, on a tellement insisté à rester dans le cadre DIFF. Bien entendu, on travaille, en fait, en dimension $n = 4$ où entre TOP et DIFF il y a un immense gouffre. Mais je serai encore beaucoup plus précis. A plusieurs reprises (dans les théorèmes C, F,...) les démarches du programme nécessitent de pouvoir appliquer des arguments du type *Hauptvermutung*, disons à partir de deux décompositions cellulaires d'un même objet, passer à une sous-division commune, mais ceci en dimension $n = 4$ (et non pas $n = 3$). Il n'est pas inutile ici de faire un bref passage en revu de cette question. En dimension $n = 3$ (ou $n \leq 3$) le Hauptvermutung est vrai dans tous les contextes possibles, résultat classique de E. Moise [54]. En dimension $n \geq 5$ D. Sullivan [62] et R. Kirby-L. Siebenmann [28] ont montré que, *modulo des obstructions* homologiques, discrètes, du type "classes caractéritiques", le Hauptvermutung est vrai dans la catégorie TOP.

En revanche, en dimension $n = 4$ (justement celle avec laquelle on travaille dans le programme) le Hauptvermutung TOP est très violemment faux, même là où aucune obstruction discrète ne peut se présenter : en utilisant des méthodes de la théorie de jauge (gauge theory) C. Taubes [65] montre que R^4 possède une infinité non-démontrable de structures PL toutes distinctes.

Mais il reste, quand même, une variante du Hauptvermutung, vraie sans aucune obstruction et en toute dimension (y compris $n = 4$) ; il s'agit du Hauptvermutung DIFF de J.H.C. Whitehead [68]. C'est le seul Hauptvermutung auquel le programme peut se raccrocher, d'où le grand intérêt à se cantonner dans la catégorie DIFF.

On va détailler maintenant, un peu plus, l'étape III du programme. On considère, pour

commencer un vrai link qu'on va concevoir ici sous la forme 4-dimensionnelle

$$\mathrm{LK} \underset{\text{def}}{=} \left(B^4_{\text{standard}}, \sum_1^k S^1_i \times \operatorname{int} D^2_i \right) \tag{4.26}$$

où $X^3 = \sum\limits_1^k S^1_i \times \operatorname{int} D^2_i \subset \partial B^4$. Un tel LK est une invitation explicite pour ajouter des anses d'indice $\lambda = 2$ à B^4, pour passer à

$$B^4 \underbrace{\cup}_{\mathrm{LK}} \sum_1^k D^2_i \times D^2_i \,, \tag{4.27}$$

suivant la recette $\sum\limits_1^k \partial D^2_i \times D^2_i = X^3$. Mais, comme on l'a déjà vu, à notre LK (4.26) s'associe, aussi, l'espèce de link V^4 du (4.9) ; on pourra remarquer qu'ici, le très classique groupe fondamental (du complémentaire) du link, $\pi_1(S^3 - X^3)$ est isomorphe au groupe fondamental à l'infini de V^4, le $\pi_1^\infty V^4$.

Le point important pour nous, c'est que ce sont justement ces π_1^∞, a priori plus ésotériques, sur lesquels notre technique, qu'on évoquera ci-dessous, à une prise ; c'est par ce biais qu'on contrôlera le groupe classique du link.

Maintenant, si on ajoute au B^4 du (4.26) un terme $\# \infty \# (S^2 \times D^2)$, somme connexe le long du bord, loin de X^3, alors on obtient une variété lisse, 4-dimensionnelle, non-compacte et à bord non-vide, rattachée canoniquement au LK

$$W^4(\mathrm{LK}) = (B^4 \# \infty \# (S^2 \times D^2)) - \{ \underbrace{\partial(B^4 \# \infty \# (S^2 \times D^2))}_{\substack{\text{ceci est la somme connexe ordinaire (entre variétés sans bord)} \\ S^3 \# \infty \# (S^2 \times S^1)}} - X^3 \} \,. \tag{4.28}$$

On a, manifestement,

$$\partial W^4(\mathrm{LK}) = X^3 \,. \tag{4.29}$$

Tout ceci ayant été dit, le résultat technique principal de l'étape III, et qui implique le Théorème C est le THÉORÈME DE LA COMPACTIFICATION ÉTRANGE (the strange compactification theorem) que voici

Théorème G. *Soit Δ^3 une boule d'homotopie telle que $\Delta^3 \times I$ est g.s.c. Il existe alors un link LK* (4.26) *avec les deux propriétés suivantes.*

1) *On a un difféomorphisme (entre variétés 4-dimensionnelles compactes, à bord non vide)*

$$(\Delta^3 \times I) \# p \# (S^2 \times D^2) = B^4 \underbrace{\cup}_{\mathrm{LK}} \sum_1^p D^2_i \times D^2_i \,. \tag{4.30}$$

2) *Il existe une variété* 4*-dimensionnelle lisse* W^4*, non-compacte à bord connexe et simplement connexe* $W^3 = \partial W^4$*, qui est munie d'un plongement lisse*

$$(W^4(\mathrm{LK}), \underbrace{\partial W^4(\mathrm{LK})}_{X^3}) \xrightarrow{\xi} (W^4, W^3) \tag{4.31}$$

tel que
2-i) $\xi \mid \operatorname{int} W^4(\mathrm{LK})$ *est un difféomorphisme*

$$\operatorname{int} W^4(\mathrm{LK}) \approx \operatorname{int} W^4$$

2-ii) $\pi_1(W^3 - \xi X^3) = F_p$ *(le groupe libre à* p *générateurs).*

Il faut penser ici à $W^4(\mathrm{LK})$ comme possédant *deux compactifications*, celle "standard" qui est $(B^4 \# \infty \# (S^2 \times D^2)) \cup \{\infty\}$ (voir (4.28)) mais aussi, la compactification *étrange* $W^4 \cup \{\infty\}$, obtenue à partir du (4.31) du théorème ci-dessus. Chacune de ces compactifications (qui sont très différentes entre elles !) donne lieu à un calcul du $\pi_1^\infty W^4(\mathrm{LK})$. Ainsi on trouve

$$\pi_1^\infty W^4(\mathrm{LK}) = \varprojlim F_n, \text{ à partir de la compactification étrange,}$$

mais aussi

$$\pi_1^\infty W^4(\mathrm{LK}) = \varprojlim (\pi_1(S^3 - X^3) * F_n), \text{ à partir de la compactification standard.}$$

Les deux $\varprojlim$ apparaissant ici sont des objets énormes mais, par des manipulations standard, leur égalité implique que le groupe $\pi_1(S^3 - X^3)$ est libre. Un théorème classique de Papakyriakopoulos nous dit alors que LK est trivial et donc l'égalité (4.30) du théorème G implique (sous notre hypothèse que $\Delta^3 \times I$ est g.s.c.) le (4.3) qui, comme on l'a déjà remarqué, donne $\Delta^3 = B^3$.

Quelques détails supplémentaires concernant le théorème de la compactification étrange

La paire (W^4, W^3) est construite par un processus infini du même type que celui déjà utilisé pour le théorème de la non-sauvagerie lisse, mais en partant maintenant d'un Δ^3 tel que $\Delta^3 \times I$ est g.s.c. Grossomodo, le LK du théorème G apparaît à un cran fini, mais déjà très haut du processus en question, ce qui en fait un objet pas du tout explicite (on ne peut pas "le dessiner"). Mais comme W^3 lui-même est le produit final d'un processus infini, c'est une 3-variété ouverte simplement connexe qui est *sauvage*, dans le sens que ses $\pi_2 W^3$ et $\pi_1^\infty W^3$ sont, tous les deux, très gros. Par ailleurs, le même processus infini, crée automatiquement deux structures supplémentaires que voici.

S_1) Soit $\Gamma = \sum_1^p S_i^1 \times 0$ le cœur (= spine) de X^3. Via ξ (4.31), on a le link "classique" $\Gamma \subset W^3$. On a alors un plongement PROPRE.

$$\sum_1^p (D_i^2 - \{\text{un Cantor non-sauvage de int } D_i^2\}) \underset{\Phi}{\hookrightarrow} W^3\,, \tag{4.32}$$

qui étend notre link classique $\Gamma \subset W^3$. [$n \geq 2$, un plongement Cantor $\subset M^n$ est dit *non-sauvage* s'il se factorise par un arc lisse $A \subset M^n$.] Maintenant, si W^3 était simplement connexe à l'infini (ce qui n'est pas le cas) alors, cette manière (4.32) de "tirer Γ à l'infini de W^3", impliquerait que le groupe $\pi_1(W^3 - \Gamma)$ est libre, et notre affaire serait finie.

S_2) Le processus infini crée aussi une *lamination* $\mathcal{L}$ de W^3, dont toutes les familles sont des R^2. [Je rappelle que les laminations $\mathcal{L}$ sont des objets analogues aux feuilletages $\mathcal{F}$. Un $\mathcal{L}$ est donné par un fermé $K=K(\mathcal{L})\subset W^3$, tel que W^3 possède un atlas $W^3=\bigcup_i u_i$ avec

$$(u_i, u_i\cap K)=(R^2\times R,\ R^2\times(\text{Cantor}\subset R))\ .]$$

Notre $\mathcal{L}$ possède, aussi, les deux propriétés suivantes (qui, entre parenthèse soit-il dit, demandent beaucoup de travail dans la conception du processus infini)

(4.33.1) $K\cap\Gamma=\emptyset$,

(4.33.2) Chaque composante connexe $U^3\subset W^3-K$ est simplement connexe à l'infini.

A l'intérieur de la réunion finie des U^3 qui le contiennent, Γ est encore tiré à l'infini, comme dans (4.32) et, via (4.33.2) on en déduit que le link classique $\Gamma\subset W^3-K$ est non-noué et non-enlacé (trivial), et donc $\Gamma\subset W^3$, aussi.

Avec ceci on a fini tout ce qu'on pouvait raisonnablement raconter sur le programme Poincaré , dans cet article. On aura remarqué le grand rôle que les pro-groupes π_1^∞ y jouent. D'habitude, les mathématiciens qui travaillent sur l'Hypothèse de Poincaré ont le souci que leurs démarches ne "prouvent" pas que toute sphère d'*homologie* soit un S^3. La structure de mon programme fait que dès le début ce type particulier de danger est écarté ; on commence dans [36], [37] par un théorème de "représentation" qui représente, justement, tous les Σ^3 et rien d'autre. En revanche, mon souci constant, au moins pendant l'étape III (mais aussi avant) a été de ne pas "prouver" que $\text{Wh}^3 = R^3$.

Vu ce rôle important des π_1^∞, il se trouve qu'il y a, aussi, une retombée importante du programme, concernant les $\pi_1^\infty\widetilde{M}^3$, question dont traitera la section V qui suit.

Quelques autres commentaires

Les espèces de links V^4 qui sont apparues dans le programme Poincaré sont les cousins beaucoup plus sophistiqués des nœuds et links tant étudiés aujourd'hui. Pour moi, parmi les questions nouvelles qu'ils soulèvent est celle de leur sauvagerie ou de leur non-sauvagerie lisse ; on a vu déjà, et ça souligne l'enjeu du problème, que l'existence d'anses de Casson différentiablement sauvage est très intimement liée à l'existence des R^4 exotiques. Ceci soulève le problème de trouver des moyens pour détecter la non-sauvagerie lisse. Les méthodes de la topologie classique (algébrique ou autre) sont absolument inopérantes ici. Mon espoir est qu'on pourra construire des *théories topologiques quantiques des champs* (topological quantum field theories) ou TQFT appropriées pour détecter des bouts différentiablement sauvages, en dimension $n=4$.

Bien entendu, les TQFT [3], [70], [71] ont fait leurs preuves en dimensions $n=3$, en particulier en théorie des nœuds, justement L. Funar a montré [20] qu'elles peuvent détecter des bouts sauvages (dans le sens de $\pi_1^\infty\neq 0$), pour des 3-variétés ouvertes. Son travail est, je crois, une sorte de "toy-model" pour les TQFT que j'ai en tête, pour la dimension $n=4$. Il y a des axiomes nouveaux (par rapport, disons, à ceux de [3]) que de telles théories, que personne ne sait construire aujourd'hui, devraient satisfaire. Il y a quelques années, j'ai longuement parlé là-dessus avec M. Martellini, un physicien italien, passé mathématicien dans la foulée de Witten. Il m'a expliqué que si j'avais les TQFT désirées, je pourrai en tirer, aussi, des renseignements sur la théorie quantique des trous noirs.

5 La conjecture d'uniformisation topologique et la simple connexité à l'infini des revêtements universels (en dimension $n = 3$)

Il y a une conjecture moins riche que la géométrisation (conjecturale, elle aussi) de Thurston, mais plus forte que la Conjecture de Poincaré . On considère maintenant une

3-variété fermée (compacte à bord vide) arbitraire, M^3 et l'espace des bouts $e(\pi_1 M^3)$ (the space of ends) de son $\pi_1 M^3$. On rappelle que, suivant H. Hopf on a seulement les quatre possibilités suivantes : $e(\pi_1(M^3)) = \emptyset$ ($\Leftrightarrow$ $\pi_1 M^3$ fini), $e(\pi_1 M^3)$ = un ou deux points, et $e(\pi_1 M^3)$ = Cantor. Par ailleurs, un très spectaculaire théorème de S. Stallings [62], dont Gromov [23] a donné une démonstration particulièrement courte, modulo les surfaces minimales, décrit la structure des groupes avec deux ou une infinité de bouts. Ceci dit, voici la

Conjecture d'uniformisation topologique. *On a*

$$\widetilde{M}^3 = S^3 - e\,(\pi_1 M^3) \tag{5.1}$$

où, le plongement $e\,(\pi_1 M^3) \subset S^3$*, évident pour* $e\,(\pi_1 M^3)$ *fini, est non-sauvage si* $e\,(\pi_1 M^3) =$ Cantor.

Cette conjecture est impliquée par la conjecture de géométrisation (ou d'uniformisation *géométrique*) de Thurston mais il n'est pas complètement déraisonnable de penser, aussi, qu'elle pourrait être une étape (autrement importante, dans ce cas) vers cette géométrisation. Par ailleurs, la conjecture d'uniformisation topologique est exactement la juxtaposition de la conjecture de Poincaré (qui, disons, décrirait le $\widetilde{M}^3$ loin de l'infini) et du théorème suivant, qui décrit le $\widetilde{M}^3$ au voisinage de l'infini, et dont je n'ai achevé la preuve que très récemment [45] (voir aussi [49]).

Théorème H. *Soit* M^3 *une variété fermée. Alors, le revêtement universel* $\widetilde{M}^3$ *est simplement connexe à l'infini.*

Remarque notationnelle importante. Dans la présente section on ne parlera jamais du π_1^∞ en tant que (pro)-groupe, et on adoptera la convention suivante, pour tout espace localement compact X : "$\pi_1^\infty X = 0$" (respectivement "$\pi_1^\infty X \neq 0$") veut dire que X est (respectivement n'est pas) simplement connexe à l'infini. Ainsi, le théorème H, sujet de la présente section, s'énonce, tout simplement comme $\pi_1^\infty \widetilde{M}^3 = 0$. Pour être encore plus spécifique, ce que la démarche menant au théorème H nous fournit, est la chose suivante. Pour tout compact $K \subset \widetilde{M}^3$ on trouvera une sous-variété N^3 de $\widetilde{M}^3$, *compacte et simplement connexe*, avec

$$K \subset N^3 \subset \widetilde{M}^3\,, \tag{5.2}$$

propriété que le lecteur pourra comparer à la s.c.g. à longue distance (4.19), s'il (ou elle) désire le faire. Mais, ce qui importe est que, vu qu'on est en dimension $n = 3$, ce qui fait que le ∂N^3 est une collection de 2-sphères et donc, via Van Kampen, (5.2) implique $\pi_1^\infty \widetilde{M}^3 = 0$. Plus généralement parlant si V^3 est ouverte et admet une exhaustion par des (polyhèdres) compacts simplement connexes, alors $\pi_1^\infty V^3 = 0$; ceci est *faux* en toute dimension $n > 3$.

Le théorème H possède, aussi, l'énoncé équivalent que voici.

Théorème H'. *Soit* M^3 *fermée, irréductible, avec* $\pi_1 M^3$ *infini, alors,* $\widetilde{M}^3 = R^3$. ["Irréductible" veut dire que tout $S^2 \subset M^3$ borde un $B^3 \subset M^3$.]

Personnellement, je préfère la formulation H, car la preuve n'a (presque) rien à voir avec le monde des variétés 3-dimensionnelles. Je pense au théorème $\pi_1^\infty \widetilde{M}^3 = 0$ comme étant, plutôt, un théorème de théorie des groupes déguisé en énoncé topologique. En fait il n'est

pas difficile de donner un sens à "$\pi_1^\infty G$" ($= 0$ ou $\neq 0$) pour un groupe de présentation finie quelconque ; alors, le théorème H dit que pour un G qui se présente comme $G = \pi_1 M^3$ on a $\pi_1^\infty G = 0$. Du point de vue de la théorie des groupes à la Gromov (ou même de l'algèbre homologique, tout court) le $\pi_1^\infty G$ est quelque chose sur lequel beaucoup reste encore à dire.

Avant d'aller plus loin, voici un très court historique. Il est très facile de voir que le théorème H est, encore, équivalent à l'énoncé suivant : *Si M^3 est fermée et $M^3 = K(\pi, 1)$, alors $\pi_1^\infty \widetilde{M}^3 = 0$.*

Il y a quelques trente ans S. Novikov conjecturait que pour toute variété M^n fermée, avec $M^n = K(\pi, 1)$, on avait $\pi_1^\infty \widetilde{M}^n = 0$. Ici je ne me base que sur mes propres souvenirs, mais je crois que ça se passait à la même époque (et dans le même contexte) où Novikov avançait sa fameuse Conjecture des hautes signatures.

Indépendamment de tout ceci, vers 1964 Waldhausen prouvait que si M^3 est de Haken (la même catégorie des 3-variétés pour lesquelles, bien plus tard (et avec certaines restrictions), Thurston [66] a prouvé son très spectaculaire théorème d'hyperbolisation), alors $\widetilde{M}^3 = R^3$.

Mais, en 1983 M. Davis [9] montre que pour toute dimension $n \geq 4$ il existe des $M^n = K(\pi, 1)$, fermées et avec $\pi_1^\infty \widetilde{M}^n \neq 0$. Comme commentaire, très a posteriori, les $\pi_1 M^n$ de Davis étaient, en fait, hyperboliques de Gromov [24] (voir [10]).

Le seul cas ouvert de la Conjecture (de Novikov) restait ainsi la dimension $n = 3$. Les premiers progrès sur la question sont obtenus, il y a environ 10 ans, indépendamment par Po [47], [48], A. Casson [63], Bestvina-Mess [4] (et quelques autres, quelques temps après ; voir aussi [53], [54]). La forme générale de tous ces résultats est l'implication suivante

$$\text{Si } \pi_1 M^3 \text{ a une géométrie "gentille"} \Rightarrow \pi_1^\infty \widetilde{M}^3 = 0. \tag{5.3}$$

Le mot "gentil" couvre ici une longue liste de propriétés concernant le graphe de Cayley de $\pi_1 M^3$, conçu comme espace métrique : hyperbolicité de Gromov, presque convexité (almost convex) de J. Cannon [5], propriété plus générale que celle de Gromov, peignabilité (combable) de Thurston, qui généralise les groupes automatiques [12], condition(s) de Casson [63], e.a.d.s. On expliquera, plus loin, la raison plus profonde qui se cache derrière (5.3). Par ailleurs, vers la même époque, j'ai prouvé le résultat suivant [46]

Théorème I. *Soit V^3 une variété ouverte simplement connexe, qui possède la propriété que pour un certain $n \in Z_+$ il existe une sous-variété lisse, PROPRE et g.s.c. $X^{n+3} \subset V^3 \times B^n$, telle que*

$$V^3 \times 0 \subset X^{n+3} \subset V^3 \times B^n\,. \tag{5.4}$$

Alors $\pi_1^\infty V^3 = 0$.

Il faut dire ici tout de suite que cette connexion, inattendue à l'époque, entre π_1^∞ et g.s.c. ne fonctionne qu'en dimension $n = 3$. [Par ailleurs, pour être précis, [46] ne prouve que le cas $X^{n+3} = V^3 \times B^n$ du théorème, mais les techniques s'adaptent facilement pour le cas général. Enfin, depuis, L. Funar [21] et L. Funar-T. Thickstun [22] ont considérablement généralisé le théorème I.]

Ceci dit, le théorème I n'est qu'un tout premier petit pas dans la direction menant au théorème H. Mais le théorème I signale, aussi, que la démarche menant à $\pi_1^\infty \widetilde{M}^3 = 0$ est

liée au (ou inspirée par) le programme Poincaré : dans les deux cas on passe de $n = 3$ à des dimensions supérieures et la simple connexité géométrique (dans *ces* dimensions) y joue un rôle fondamental. Mais, tandis que pour le programme Poincaré , il faut monter à la dimension $n = 4$ et pas plus, maintenant on peut se permettre de monter dans des dimensions très grandes, insensibles aux difficultés particulières de la dimension $n = 4$ (ou $n = 3$). Mais dans les deux programmes, il y a des énormes difficultés "combinatoires" à affronter. Dans le cadre des étapes I et III du programme Poincaré, ces difficultés ont été complètement repoussées vers l'infini (corraled at infinity), par les processus infinis, justement ; c'est encore avec de tels processus infinis, accompagnés de compactifications adaptées, qu'avec D. Gabai on est en train de s'attaquer à la Conjecture B. Tout ceci ayant été dit, l'obstruction précise qu'on doit affronter pour le $\pi_1^\infty \widetilde{M}^3$ est que les difficultés combinatoires ne se laissent *pas* entièrement repousser à l'infini (sauf si $\pi_1 M^3$ est à géométrie gentille, comme dans (5.3)) ; une partie *résiduelle*, le "*mauvais locus*", dont il sera question plus loin survit à distance finie, dans le cas général, et doit être affronté.

Le premier pas sérieux dans la preuve de $\pi_1^\infty \widetilde{M}^3 = 0$ est le THÉORÈME DE REPRÉSENTATION ÉQUIVARIANTE DE $\widetilde{M}^3$ [52] qui, en s'inspirant d'un pas analogue mais beaucoup plus modeste du tout début de la preuve du théorème D (un petit pas développé dans [36]) construit un objet

$$X^2 \xrightarrow{f} \widetilde{M}^3 \tag{5.5}$$

avec une longue liste de propriétés qu'en une première approximation je résume ainsi. Pour commencer, X^2 est un complexe simplicial localement fini qui est g.s.c. (dans le sens que tous ses voisinages réguliers $N^p(X^2)$, où p est assez grand, le sont) f est non-dégénérée et, en dehors d'un ensemble de singularités $\mathrm{Sing}(f) \subset X^2$ c'est une immersion générique, l'ensemble des points doubles $M^2(f) \subset X^2 \times X^2$ peut être épuisé en "tirant la fermeture éclair à partir des singularités de f" (f is zippable) et fX^2 remplit "essentiellement" $\widetilde{M}^3$. La liste (très approximative et schématique) des propriétés de (5.5) ne s'arrête pas là, mais avant de continuer on doit faire une parenthèse. Maintenant, il y a, aussi l'ensemble des points doubles "en bas" $M_2(f) \subset X^2$ (image du $M^2(f)$ par $X^2 \times X^2 \to X^2$, projection sur le premier facteur) et si $M_2(f)$ (complété par $\mathrm{Sing}(f)$) était *fermé* dans X^2 on pourrait, en fait, conclure à partir de tout ce que je viens déjà de dire que $\pi_1^\infty \widetilde{M}^3 = 0$. Mais les choses ne peuvent pas être aussi simples que ça, car le thórème de représentation que je viens d'esquisser s'applique, en fait, à *toute* variété V^3 ouverte et simplement connexe, y compris à $V^3 = \mathrm{Wh}^3$ (4.15). Ainsi, $M_2(f)$ n'est PAS fermé, généralement parlant (donc f n'est pas une application simpliciale non plus) et on a un *mauvais locus* constitué essentiellement, par les points d'accumulation indésirables de $M_2(f)$; en présence des points triples $M_3(f)$ c'est un peu plus compliqué que $\overline{M_2(f)} - M_2(f)$ auquel on pourrait penser. [Remarque. Si M^3 est tel que $\pi_1 M^3$ a une géométrie gentille (5.3), alors on peut construire une représentation $X^2 \xrightarrow{f} \widetilde{M}^3$ avec les propriétés qu'on vient déjà de passer en revue et, aussi, avec

$$\{\text{mauvais locus}\} = \emptyset\,.$$

C'est ceci qui se cache derrière (5.3), si on creuse assez loin.]

Une fois tout ceci ayant été dit, notre (5.5) du théorème de représentation équivariante, possède encore les deux propriétés suivantes.

i) Le mauvais locus est bien là, mais on peut le contrôler. Soit $\Lambda \subset X^2$ une *transversale tendue* et pré-compacte à $M_2(f)$; on considère l'ensemble des points d'accumulation de $\Lambda \cap M_2(f) \subset \Lambda$, qu'on notera $\lim(\Lambda \cap M_2(f))$. Avec ceci

$$\text{card}\,(\Lambda \cap M_2(f)) < \infty\,. \tag{5.6}$$

)em[ii) On a équivariance ; le $\pi_1 M^3$ agit librement sur X^2 et, pour $x \in X^2$, $g \in \pi_1 M^3$, on a

$$f(gx) = g\,f(x)\,. \tag{5.7}$$

Commentaire. Les deux dernières propriétés "coutent très cher". Par ailleurs si on essaie la même démarche pour la variété de Whitehead, en prenant un $X^2 \xrightarrow{f} \text{Wh}^3$ le "plus naturel possible" alors, pour des transversales tendues Λ bien choisies, le $\lim(\Lambda \cap M_2(f))$ est un Cantor, en fait, dans un certain sens, c'est un ensemble de Julia de la dynamique complexe (voir [55] pour les détails).

Quelques détails supplémentaires concernant la preuve du théorème H ($\pi_1^\infty \widetilde{M}^3 = 0$)

On a intérêt à passer de V^3 a un objet

$$V_h^3 \underset{\text{def}}{=} V^3 - \{\text{un fermé totalement discontinu et non-sauvage}\} \tag{5.8}$$

pour lequel il n'est pas difficile de voir que $\pi_1^\infty V_h^3 = \pi_1^\infty V^3$. Ce V^3 avec "beaucoup de petits trous" (V^3 with many little holes) est plus facile à contrôler, pour nous, que V^3. Pour être plus précis, dans le théorème de représentation équivariant, fX^2 n'est pas localement fini, mais le voisinage régulier *ouvert* $Nbd(fX^2) \subset \widetilde{M}^3$ est bien défini et homéomorphe à $\widetilde{M}_h^3$. Quand, dans le contexte de Wh^3 on tombe sur les ensembles de Julia, on a un f tel que $Nbd(fX^2) = Wh^3$ et non un Wh^3 avec beaucoup de petits trous. Par ailleurs, le $\widetilde{M}_h^3$, à la place de $\widetilde{M}^3$, est le prix à payer pour avoir (5.6).

Dans notre théorème de représentation équivariante, la propriété que X^2 soit g.s.c. est liée au fait que la variété représentée est ouverte et simplement connexe tandis que (5.7) est manifestement liée aux symmétries spécifiques de $\widetilde{M}^3$. Mais toutes les *autres* propriétés du théorème de représentation gardent leur sens pour, disons, une flèche $X^2 \xrightarrow{f} W^3$ où W^3 est maintenant une 3-variété arbitraire, avec $\partial W^3 = \emptyset$. On va prendre un tel $X^2 \xrightarrow{f} W^3$, pour le moment hypothétique, (et, aussi, sans équivariance et sans g.s.c) comme étant la *définition* d'une "S_u*-structure*". Le mot structure utilisé ici n'est pas anodin. Il faut penser à $X^2 \xrightarrow{f} W^3$ (qu'on notera $S_u W^3$) comme étant une masse de matériel 2-dimensionnel entassé (piled up) dans W^3 suivant des règles du jeu assez précises (toutes en fait locales sauf le "zippable"). Ceci n'est pas complètement dissemblable à un feuilletage $\mathcal{F}$ ou une limination $\mathcal{L}$; en fait, sous certaines conditions (les mêmes que pour $\mathcal{F}$ ou $\mathcal{L}$) on peut tirer S_u en arrière (pull-back) ou le pousser en avant (push-forward) en prenant son image directe, via des applications entre 3-variétés. Par exemple, (5.7) se traduit comme $g_* S_u = S_u$ ou $g^* S_u = S_u$ pour le $S_u \widetilde{M}^3$ de (5.5). Pour $n+3$ assez grand ($n+3 \geq 8$), on peut, aussi, associer canoniquement à $S_u W^3$ deux variétés lisses, non-compactes, à bord non-vide, de dimension $n+3$, on les appellera $\|S_u W^3\|$ et $|S_b W^3|$ avec les propriétés suivantes.

(5.9) Grosso-modo, $\|S_u W^3\|$ est une sous-variété PROPRE de $W_h^3 \times B^n$ telle que

$$W_h^3 \times 0 \subset \|S_u W^3\| \subset W_h^3 \times B^n$$

(5.10) *Si* W^3 est ouverte, avec $\pi_1 W^3 = 0$, alors $|S_b W^3|$ est g.s.c.

(5.11) En fait la différence entre $\|S_u W^3\|$ et $|S_b W^3|$ est très subtile. La construction de $|S_b W^3|$ "imite" celle de $\|S_u W^3\|$ et, dès que {mauvais locus}$=\emptyset$ ($\Leftrightarrow M_2(f)$ fermé), on a $\|S_u W^3\| \underset{\text{DIFF}}{=} |S_b W^3|$. Notons que si

$\|S_u W^3\| \underset{\text{DIFF}}{=} |S_b W^3|$ alors, en combinant (5.9), (5.10) et le théorème I, on tire que $\pi_1^\infty W^3 = 0$. Comme W^3 pourrait être ici la variété de Whitehead Wh^3, on voit bien qu'on ne peut pas toujours avoir $\|S_u W^3\| = |S_b W^3|$.

Avant d'aller plus loin, je fais quelques remarques. Je dis qu'on va canoniquement de $S_u W^3$ à $\|S_u W^3\|$ et $|S_b W^3|$. Le mot "canonique" ici inclut le fait que S_u, S_b héritent de toutes les symétries de $X^2 \xrightarrow{f} W^3$. Par ailleurs, la construction d'un $|S_b W^3|$ qui est g.s.c. et qui "ressemble " beaucoup à $\|S_u W^3\|$ est très délicate ; elle serait impossible sans la propriété de finitude (5.6). Le $|S_b W^3|$ provient lui aussi d'une $S_b W^3$-structure, analogue à $S_u W^3$, mais plus bizarre. On a vu, plus haut que $\|S_u W^3\| \underset{\text{DIFF}}{=} |S_b W^3|$ entraîne $\pi_1^\infty W^3 = 0$ mais, de fait, les deux propriétés sont, essentiellement, équivalentes.

Avec tout ceci on considère maintenant la structure $S_u \widetilde{M}^3$ *équivariante* et la projection $\widetilde{M}^3 \xrightarrow{\pi} M^3$. La $S_u \widetilde{M}^3$ peut être quotienté par $\pi_1 M^3$ donnant lieu à une S_u-structure

$$S_u M^3 = \pi_* S_u \widetilde{M}^3 \tag{5.12}$$

en bas. De même $S_b M^3 = \pi_* S_b \widetilde{M}^3$, $\pi^* S_u M^3 = S_u \widetilde{M}^3$, e.a.d.s. On a maintenant les deux variétés lisses de dimension $(n+3)$, $\|S_u W^3\|$ et $|S_b M^3|$, avec le même π_1 que M^3. Ce qu'on a gagné, c'est que M^3 est *compact* (et compact $\Rightarrow \pi_1^\infty = 0$). Les choses sont un peu plus compliquées que ça, mais (avec pas mal de travail) de la compacité de M^3, on peut tirer un difféomorphisme

$$\|S_u M^3\| = |S_b M^3| \,. \tag{5.13}$$

Le (5.13) implique, manifestement, que

$$\|S_u M^3\| = |S_b M^3| \quad . \tag{5.14}$$

Toute notre petite théorie est fonctorielle, ce qui fait que $\|S_u M^3\| = \|S_u \widetilde{M}^3\|$ et $|S_b M^3| = |S_b \widetilde{M}^3|$. Ainsi $\|S_u \widetilde{M}^3\| = |S_b \widetilde{M}^3|$; d'après (5.11) ceci entraîne que $\pi_1^\infty \widetilde{M}^3$. J'insiste encore une fois sur le fait que ce résumé de preuve est hautement schématisé et simplifié.

6 Mathématiques et ontologie (Un essai provisoire)

Il y a au moins deux points de vue, les deux relativement bien connus, concernant le statut existentiel des êtres mathématiques : le point de vue platonicien (plus ou moins majoritaire chez les mathématiciens) et le point de vue constructiviste ou intuitionniste, relativement minoritaire. On va se placer ici, carrément, dans la perspective platonicienne, disons, par exemple, telle qu'Alain Connes l'explique très clairement dans ses conversations avec un neuro-physiologue dont le matérialisme aurait pu faire pâlir d'envie, justement, l'auteur de "Matérialisme et Empiriocriticisme" [**?**]. Soit dit entre parenthèse, j'ai trop longtemps écouté mon ami Gabriel Stolzenberg, l'un des plus convaincants apôtres du constructivisme, pour ne pas être sensible, aussi à son point de vue, mais il sera plus simple pour nous ici d'assumer, dès le départ une vue platonicienne des mathématiques.

En fait, ce que je vais proposer comme sujet de réflexion va un peu plus loin. Une affirmation comme l'hypothèse de Poincaré concerne-t-elle, disons, l'ensemble des diagrammes de Heegard simplement connexes (et ici j'aurais pu choisir, tout aussi bien, toute autre incarnation terrestre des sphères d'homotopie) ou plutôt une propriété qui existe au niveau du concept même de sphère d'homotopie ? Quel est, par exemple, le statut existentiel des "espèces de link" évoqués dans la section IV ou du "mauvais locus" de la section V ? Existe-t-il aux niveaux extensionnels des totalités des diagrammes de Heegard (avec $\pi_1 = 0$) ou des $\widetilde{M}^3$, ou plutôt au niveau de la compréhension des concepts respectifs de sphère d'homotopie ou de revêtement universel ? De longues années passées en compagnie de ces objets me font incliner plutôt vers le point de vue "réaliste", et j'utilise ici ce mot dans le sens médiéval et

scolastique de John Scot Eurigena, plutôt que vers le point de vue "nominaliste", dans le sens tout aussi médiéval et scolastique d'Abélard.

Ceci n'est pas une simple question métaphysique, elle a une incidence directe sur la manière de travailler, tous les jours, sur les problèmes en question. Va-t-on se faire la main sur des objets concrets (disons nos fameux diagrammes de Heegard, par exemple ou de pavages en domaines fondamentaux bien beaux et symétriques) ou sur des "théories", beaucoup plus abstraites et opérant, éventuellement, sur des objets sauvages ? Je vais essayer d'être un peu plus explicite sur ce point. Disons qu'on commence par fractionner le problème très concret qu'on a en vue, en un nombre d'énoncés très ésotériques (qui, ensemble, impliquerait quand même la chose qu'on veut). Par leur même nature, ces énoncés ésotériques ne se prêtent plus à des manipulations concrètes sur des objets mathématiques concrets. On ne peut, alors, que fabriquer des "théories" qui, on l'espère, vont prouver ce que l'on veut prouver. Souvent, ces théories se laissent tester sur des objets monstrueusement sauvages, comme je l'ai dit plus haut. Même quand il ne s'agit pas de monstruosités tout court, pour l'Hypothèse de Poincaré mon test n'a jamais été le classique "s'assurer qu'on n'a pas prouvé que toute sphère d'homologie est S^3" mais plutôt "s'assurer qu'on n'a pas prouvé que la variété de Whitehead est simplement connexe à l'infini." D'habitude aussi, ces théories procèdent en deux étapes distinctes. Il y a une première étape formelle, un processus infini peut-être heuristique, qui, à l'aide peut-être d'une manipulation légèrement paradoxale, prouverait ce qu'on veut démontrer ; en fait, il ne le fait pas, car il ne converge pas. Vient alors la seconde étape, "la renormalisation" qui met en place un second processus infini qui imite (ou essaye d'imiter) le processus formel mais qui, cette fois-ci converge. Même quand ceci fonctionne, on ne démontre, finalement, qu'un énoncé plus faible que celui auquel on aurait pu espérer. D'où la nécessité de fragmenter le problème initial en *plusieurs* énoncés de caractère ésotérique. Mais la plupart des théories ne sont pas renormalisables ou, même quand elles le sont on ne se retouve, à la fin, qu'exactement au point de départ (en ayant, peut-être, perdu une année de travail).

J'espère que j'ai assez dit pour donner l'idée de comment le point de vue réaliste (plutôt que celui nominaliste) peut être un peu plus qu'une vue purement philosophique, et avoir une incidence pratique directe, être en un certain sens "testé".

Je termine avec un petit souvenir qui n'est peut-être pas tout à fait déplacé, au point où on est arrivé. En 1979, j'étais à une conférence sur les théories de jauge à Cargèse où l'une des vedettes était une étoile alors montante, dont on a beaucoup entendu parler depuis, Ed Witten. Il se trouve qu'on était dans le même hôtel et que j'ai eu donc souvent l'occasion de l'écouter ; et il y a une chose qui m'a beaucoup frappé. Tandis que les autres physiciens que je connaissais parlaient toujours en termes de phénomènes, éventuellement en termes de modèles concrets testés sur ordinateurs, Witten jonglait tout le temps avec les théories elles-mêmes, leurs manières d'interagir, de se compléter ou tout simplement d'exploser. Il me semble qu'il y a un lien direct entre cette manière de penser et la M-théorie d'aujourd'hui. Mais ceci est une autre histoire.

Bibliographie

[1] J.W. Alexander, *On the subdivision of 3-space by a polyhedron*, Pr. N. A. S. **10**, 1924, pp. 6-8.

[2] E. Artin, R.H. Fox, *Some wild cells and spheres in 3-dimensional space*, Ann. of Math. **49**, 1948, pp. 979-990.

[3] M. Atyiah, *Topological quantum field theories*, Publ. IHES **68**, 1989, pp. 175-186.

[4] M. Bestvina, G. Mess, *The boundary of negatively curved groups*, J. Ann. Math. Soc. **4** no. 3, 1991, pp. 469-481.

[5] J.W. Cannon, *Almost convex groups*, Geometria Dedicata **22**, 1987, pp. 197-210.

[6] A. Casson, *Three lectures on new-infinite constructions in 4-dimensional manifolds*, dans le volume *A la Recherche de la Topologie Perdue*, L. Guillou, A. Marin, (edts.) Birkhaüser, 1986, pp. 201-244.

[7] J. Cerf, *Sur les difféomorphismes de la sphère de dimension trois ($\Gamma_4 = 0$)*, SLN **53**, 1968.

[8] J.P. Changeux, A. Connes, *Matière à Pensée*, O. Jacob, 1989.

[9] M.W. Davis, *Groups generated by reflexions and aspherical manifolds not covered by Euclidean Spaces*, Ann. of Math. **117**, 1983, pp. 293-324.

[10] M. Davis, M. Shapiro, *Coxeter groups are almost convex*, Preprint, 1991.

[11] S. Donaldson, *An application of gauge theory to the topology of 4-manifolds*, J. Diff. Geom. **18**, 1983, pp. 279-315.

[12] D.B.A. Epstein, J.W. Cannon, D.F. Holt, S.V.F. Levy, M.S. Patterson, W. Thurston, *Word processing in groups*, James en Barlett 1992.

[13] D. Gabai, Valentin Poénaru's Program for the Poincaré Conjecture, in the volume *Geometry Topology and Physics for Raoul Bott* (ed. by S.T. Yau) International Press , 1994, pp. 139-169.

[14] M. Gerstenhaber, O. Rotham, *The solution of sets of equations in groups*, Proc. Nat. Acad. Sci. **48** , 1962, pp. 1531-1533.

[15] R.E. Gompf, *Three exotic R^4, and other anomalies*, J. Diff. Geom. **18**, 1983, pp. 317-328.

[16] E. Ghys, P. de la Harpe, (editors) *Sur les groupes hyperboliques de Mikhael Gromov*, Birkhäuser, 1990.

[17] M. Freedman, *The topology of four-dimensional manifolds*, J. Diff. Geom. **17**, 1982, pp. 357-453.

[18] M. Freedman, *There is no room to spare in four-dimensional space*, *Notices AMS* **31**, 1984, pp. 3-6.

[19] M. Freedman, F. Quinn, *Topology of 4-manifolds*, Princeton University Press, 1990.

[20] L. Funar, *TQFT and the Whitehead manifold*, J. Knot Th. **6**, 1997, pp. 13-30.

[21] L. Funar, *Infinite simple homotopy and the simple connectivity at infinity of open 3-manifolds*, Preprint Université de Grenoble, 1996, see also : On proper homotopy type and simple connectivity at infinity for open 3-manifolds, Preprint, 1999.

[22] L. Funar, T. Thickstun, *On open 3-manifolds proper homotopy equivalent to geometrically simply connected polyhedra*, Preprint Grenoble, 1998.

[23] M. Gromov, *Infinite groups as geometric objects*, Proc. ICM Warszawa, 1983, pp. 385-392.

[24] M. Gromov, *Hyperbolic groups*, in *Essays in Group Theory* (S.M. Gerstein, ed.), MSRI Publ. **8**, 1987, pp. 72-263.

[25] M. Kervaire,*A manifold which does not admit any differentiable structure*, Comment. Math. Helv. **34**, 1960, pp. 304-312.

[26] M. Kervaire, J. Milnor, *Groups of homotopy spheres*, I, Ann. of Math. **77**, 1963, pp. 504-537.

[27] R. Kirby, *The topology of 4-manifolds*, SLN **1374**, 1989.

[28] R. Kirby, L. Siebenmann,*Foundational essays on triangulations and smoothings of topological manifolds*, Ann. of Math. Studies **88**, 1977.

[29] B. Mazur,*On embeddings of spheres*, BAMS **65**, 1959, pp. 59-65.

[30] B. Mazur,*A note on some contractible 4-manifolds*, Ann. of Math.. **73**, 1961, pp. 221-228.

[31] J. Morgan, H. Bass,*The Smith Conjecture*, Academic Press, 1984.

[32] J. Milnor, *On manifolds homeomorphic to the 7-sphere*, Ann. of Math. **64**, 1956, pp. 399-405.

[33] E.E. Moise, *Certain classical problems of euclidean topology*, *Lectures of Summer Institute on Set Theoretic Topology*, Madison, 1955.

[34] R. Penrose, J.H.C. Whitehead, E.C. Zeeman, *Imbedding of manifolds in euclidean space*, Ann. of Math.. **73**, 1961, pp. 613-623.

[35] H. Poincaré, *Cinquième complément à l'Analysis Situs*, Rend. Circ. Mat. Palermo **18**, 1904, pp. 45-110.

[36] V. Poénaru, *The collapsible pseudo-spine representation theorem*, Topology, vol. 31, **3**, 1992, pp. 625-636.

[37] V. Poénaru, *Infinite processes and the 3-dimensional Poincaré Conjecture, II : The Honeycomb representation theorem*, Prépublications d'Orsay 93-14, 1994.

[38] V. Poénaru, *Infinite processes and the 3-dimensional Poincaré Conjecture, Conjecture III : The algorithm*, Prépublications d'Orsay 92-10, 1992.

[39] V. Poénaru, *Processus infini et conjecture de Poincaré en dimension trois, IV : Le théorème de non sauvagerie lisse (The smooth tameness theorem), Part A*, Prépublications d'Orsay 93-83, 1992 ; Part B, Prépublications d'Orsay 95-33, 1995.

[40] V. Poénaru, *Geometric simply connectivity in smooth four-dimensional topology*, Prépublications d'Orsay 98-63, 1998, a longer detailed version of this paper is the IHES Prépublication M/01/45, 2001, with the same title.

[41] V. Poénaru, *The strange compactification theorem, Part A*, IHES Prépublication M/95/15, 1995 ; Part B, IHES Prépublication M/96/43, 1996 ; Part C, IHES Prépublication M/97/43, 1997 ; Part D, IHES Prépublication M/97/59, 1997 ; Part E is in the process of being typed at IHES.

[42] V. Poénaru, *Les décompositions de l'hypercorde en produit topologique*, Bull. S.M.F. **88** 1960, pp. 113-129.

[43] V. Poénaru, *Three lectures on higher-dimensional topological methods in three-dimensional topology*, Suppl. Rend. Palermo, 1997, pp. 203-217.

[44] V. Poénaru, *A program for the Poincaré Conjecture, and some of its ramifications*, to appear in the volume Low-dimensional topology, Penn. State 1996 (World Scientific) A. Banyiaga, M. Movahedi-Lankarani, R. Wells, editors.

[45] V. Poénaru, *Universal convering spaces of closed 3-manifolds are simply-connected at infinity. A short outline of the proof*, Prépublications d'Orsay 98-63, 1998. A long detailed version, with the same title is Prépublications d'Orsay 2000-20, 2000. Revised versions will be out soon, for both papers.

[46] V. Poénaru, *Killing handles of index one stably and* π_1^∞, Duke Math. J. **63**, 1991, pp. 431-447.

[47] V. Poénaru, *Almost convex groups, Lipschitz Combing and* π_1^∞ *for universal covering spaces of closed 3-manifolds*, J. Diff. Geom. **35**, 1992, pp. 103-130.

[48] V. Poénaru, *Geometry "à la Gromov" for the fundamental group of a closed 3-manifold* M^3 *and the simple connectivity at infinity of* $\widetilde{M}^3$, Topology **33**, 1994, pp. 181-196.

[49] V. Poénaru, π_1^∞ *and infinite simple homotopy type in dimension three*, (H. Nenska editor) Madeira 98, Contemporary Math. **233** AMS, 1999, pp. 1-28.

[50] V. Poénaru, C. Tanasi, *Some remarks on geometric simple connectivity*, Acta Math. Hungarica **81**, 1988, pp. 1-12.

[51] V. Poénaru, C. Tanasi, *A note on the wick of the Casson Handles*, Rend. Padova **105**, 2001, pp. 1-23.

[52] V. Poénaru, C. Tanasi, *Equivariant, almost-arborescent representations of open simply-connected 3-manifolds ; A finiteness result* (second revised version, July 2001, Prépublications d'Orsay 2001-45, 2001.

[53] V. Poénaru, C. Tanasi, *Hausdorff combing of groups and* π_1^∞ *for universal covering spaces of closed 3-manifolds*, Ann. Sc. Norm. Sup. di Pisa, Serie IV, vol. XX, fasc. 3, 1993, pp. 387-414.

[54] V. Poénaru, C. Tanasi, *k-weakly almost convex groups and* $\pi_1^\infty \widetilde{M}^3$, Geometria Dedicata **48**, 1993, pp. 57-81.

[55] V. Poénaru, C. Tanasi, *Representations of the Whitehead manifold* Wh^3 *and Julia sets*, Ann. Fac. Sc. de Toulouse, Vol. IV, no. 3, 1995, pp. 655-694.

[56] V. Poénaru, C. Tanasi, *On the handles of index one of the product of an open simply-connected 3-manifold, with a high dimensional ball*, Suppl. Rendiconti Ciralo Palermo **60**, 2000, pp. 1-135.

[57] S. Smale,*Generalized Poincaré's conjecture in dimensions greater than four*, Ann. of Math. **74** no. 2, 1961, pp. 361-406.

[58] S. Smale, *On the structure of manifolds*, Ann. J. of Math. **84**, 1962, pp. 387-399.

[59] J. Stallings, *Polyhedral homotopy-spheres*, BAMS **66**, 1960, pp. 485-488.

[60] J. Stallings, *The piecewise linear structure of euclidean space*, Proc. Cambridge Phil. Soc. **58**, 1962, pp. 481-488.

[61] J. Stallings, *How not to prove the Poincaré Conjecture*, Ann. of Math. Studies **60**, 1960, pp. 83-88.

[62] J. Stallings, On torsion free groups with infinitely many ends, Ann. of Math. **88**, 1968, pp. 312-334.

[63] J. Stallings, S.M. Gersten, *Casson's idea about 3-manifolds whose universal cover is* R^3, Int. J. Algebra Comput. **1** no. 4, 1992, pp. 395-406.

[64] D. Sullivan, *On the Hauptvermutung for manifolds*, BAMS **73**, 1967, pp. 598-600 ; and *Triangulation, smoothing homotopy equivalences and homeomorphims* (Princeton Notes, 1967.

[65] C. Taubes, *Gauge theory on asymptotically periodic 4-manifolds*, J. Diff. Geom. **25**, 1987, pp. 363-430.

[66] W.P. Thurston, *Three dimensional manifolds, Kleinian groups and hyperbolic geometry*, BAMS **6**, 1982, pp. 357-381.

[67] J.H.C. Whitehead, *A certain open manifold whose groups is unity*, Q.J. of Math. **6**, 1935, pp. 268-279.

[68] J.H.C. Whitehead, *On* C^1*-complexes*, Ann. of Math. **41**, 1940, pp. 809-824.

[69] J.H.C. Whitehead, *Simplicial spaces, nuclei and m-groups*, Proc. Lond. Math. Soc. **46**, 1939, pp. 243-327 ; et aussi, Simple homotopy types, Ann. J. of Math. **72**, 1950, pp. 1-57.

[70] E. Witten, *Topological quantum field theory*, Comm. Math. Phys. **117**, 1988, pp. 353-386.

[71] E. Witten, *Quantum field theory and the Jones polynomial*, dans Braid group, knot theory and statistical mechanics C.N. Yang, M.L. Ge editors. World Scientific, 1989, pp. 239-329.

[72] E.C. Zeeman, *The generalized Poincaré Conjecture*, BAMS **67**, 1961, pp. 270.

Le retour de la géométrie

Klaus Volkert,
Frankfurt a. M. [RFA]

We give some informations about the birth of the theory of Seifert-fibered manifolds. In particular we enter into the details of the use of (spherical) geometry in two joint papers by Threlfall and Seifert (1931 and 1932) which were the first steps in that direction.[1]

1 La situation de la topologie vers la fin des années 1920

Vers la fin des annèes 1920, une certaine stagnation se fait sentir dans la topologie classique. Par topologie classique, il faut entendre " la théorie des variétés ", domaine inauguré par Poincaré avec sa fameuse série d'articles sur l'*analysis situs* (1892–1904)[2]. Le but de cette théorie était de classer les variétés aux homéomorphies près (selon le terme proposé par Poincaré ; cf. Poincaré 1953, 199) ; son paradigme était la classification des surfaces fermées[3]. L'outil principal de cette théorie était des invariants comme les nombres de Betti ou le groupe fondamental introduit par Poincaré dans la note citée de 1892. Pour définir les variétés et calculer les invariants, on utilisait souvent une structure combinatoire ; autrement dit, les variétés étaient construites à partir de composants simples (par exemple des tétraèdres généralisés). D'où la dénomination de topologie combinatoire. Cette approche était particulièrement intéressante pour aborder la question des fondements du nouveau domaine [4]. Cependant, vers la fin des années 1920, ce programme était en crise :

– Grâce aux progrès de la théorie générale de la topologie, les variétés et d'autres notions fondamentales comme les revêtements pouvaient être définies de manière abstraite sans recourir à une structure combinatoire (Hausdorff, Weyl). Par conséquent, l'existence d'une structure combinatoire apparaissait comme une restriction supplémentaire qu'il fallait justifier.

[1] Une étude plus complète et plus détaillée avec des textes inédits de Threlfall sera publiée par l'Académie des Sciences de Heidelberg. Beaucoup de gens ont contribué à ce travail. En particulier je veux remercier Ph. Nabonnand (Nancy), D. Puppe (Heidelberg), M. Romberg (Stutensee), D. Steinmetz (Karlsruhe), R. Strebel (Fribourg) et H. Struve (Köln).

[2] Bien entendu, Poincaré eut des précurseurs comme Riemann, Betti, Dyck ; cf. [30] chap.2.

[3] Cf. la note " Sur l'analysis situs " de 1892, la première publication de Poincaré consacrée à la topologie ([21], 189 – 192).

- Un point important de cette justification était la " conjecture principale " (*Hauptverrmutung*) : Deux variétés admettant une structure combinatoire homéomorphes sont-elles équivalentes du point de vue combinatoire ?
- Une autre question importante : Une variété est-elle toujours triangularisable ?
- Des difficultés énormes apparaissaient concernant les démonstrations (ou les techniques) dans le cadre combinatoire. L'exemple le plus fameux est celui du lemme de Dehn ; des erreurs, connues par Dehn lui-même depuis longtemps, dans la démonstration de lemme étaient signalées par H. Kneser [17].

En 1930 van der Waerden (travaillant lui-même à cette époque sur des questions de topologie) résumait ainsi la situation dans un discours prononcé à la réunion de la société des mathématiciens allemands (DMV) à Pragues :

> " Je pense qu'on peut constater avec une grande certitude que les difficultés posées par ces recherches sont et restent si grandes qu'il est à préférer de démontrer l'invariance des groupes d'homologie et du groupe fondamental d'une manière directe – sans le détour par la notion combinatoire d'équivalence ou d'invariance – c'est à dire par la méthode mixte. " [[29], 127]

(" Méthode mixte " veut signaler qu'on travaille avec les invariants classiques mais dans un cadre ensembliste ou continu [le continu étant opposé au combinatoire c'est à dire au discret]).

Van der Waerden était membre du groupe de mathématiciens réunis autour d'Emmy Noether, groupe qui était à l'origine de l'algébrisation de la topologie[4]. L'approche algébrique des questions topologiques peut elle aussi se comprendre comme une réaction aux difficultés de la topologie combinatoire. Si elle donnait une nouvelle orientation à la topologie en créant des domaines nouveaux comme l'algèbre homologique et la théorie des catégories, il faut constater que l'algèbrisation n'a pas beaucoup contribuée à la résolution des problèmes classiques comme la classification des variétés ! Ces problèmes ouverts furent remplacés par d'autres.

Outre la définition abstraite dans le langage de la théorie des ensembles ou la construction concrète à partir d'éléments simples, une troisième manière de construire des variétés[5] s'offrait : la voie choisie par Poincaré au début de sa série d'articles sur l'*analysis situs* qui consiste à considérer un cube dans l'espace ordinaire et certaines identifications sur les faces du cube.

Un exemple simple de cette méthode est celui de la construction du tore en dimension 3. Si les arêtes du cube ont une longueur 1, on peut utiliser les translations suivantes pour identifier les faces opposées :

$$\begin{aligned} (x,y,z) &\rightarrow (x+1,y,z) \\ (x,y,z) &\rightarrow (x,y+1,z) \\ (x,y,z) &\rightarrow (x,y,z+1) \end{aligned}$$

On peut alors se convaincre que le tore ainsi construit est localement isométrique à l'espace tridimensionnel ordinaire – un fait qui n'est pas discuté par Poincaré.

[4] Dénomination utilisée pour la première fois par Alexandroff en 1932 dans son appendice au traité de Cohn-Vossen et Hilbert, " *Anschauliche Geometrie* " de Cohn-Vossen et Hilbert.

[5] Sauf mention contraire explicite, les variétés que nous allons considérer seront tridimensionnelles, fermées et orientables.

Evidemment la géométrie joue un rôle essentiel dans cette définition qui se peut généraliser en utilisant d'autres isométries ou même des applications affines comme des transvections. Le quatrième exemple donné par Poincaré est construit en combinant les translations cités avec des rotations par 90°. Cette construction produit une variété tridimensionnelle fermée dont le groupe fondamental est fini d'ordre huit, le groupe hypercubique (le groupe multiplicatif des unités du corps des quaternions) dans la terminologie de Poincaré. Plus tard, on a nommé cette variété l'espace quaternionien ; Threlfall et Seifert ont montré que la géométrie de cette variété est la géométrie sphérique.

Je pense que la motivation de Poincaré pour étudier ces exemples (voir [31])– que l'on peut appeler " variétés cubiques " – était ses recherches, datant de la première moitié des années 1880, sur les fonctions Kleinéennes. En effet, il y avait rencontré des partitions de l'espace hyperbolique par des polyèdres réguliers en particulier par des cubes ; ces partitions étant générées par des sous-groupes discrets de $SL(2,\mathbb{C})$. Dans une langue plus moderne, la construction de Poincaré peut se décrire comme un espace homogène ou encore dans le cadre de la théorie des revêtements. Mais à l'époque de Poincaré un tel langage n'existait pas ; par conséquent Poincaré utilisait une construction tout à fait intuitive.

Poincaré n'était pas conscient que sa construction était liée à un problème dénommé par Killing, problème des formes d'espace de Clifford-Klein. Motivé par la découverte d'une surface fermée à courbure constante nulle dans l'espace elliptique (un tore) par Clifford, Klein avait proposé en 1890 le problème de déterminer toutes les surfaces à courbure constante. Ainsi formulé, il s'agissait clairement d'un problème de géométrie différentielle. Killing démontra en 1892 que la construction des formes d'espace est équivalente à la détermination de certains sous-groupes des groupes d'isométries des géométries classiques, opérant proprement discontinument et sans point fixe. D'un point de vue moderne, une forme d'espace est un espace homogène[6] ; l'espace classique étant le revêtement universel de la forme d'espace.

2 Les premiers travaux de Seifert et Threlfall

W. Threlfall (1888-1949), le plus agé de nos auteurs, était plus ou moins autodidacte en mathématiques[7]. Au début du 20^e siècle, il avait d'abord étudié la chimie pour se tourner ensuite vers les mathématiques. Il passa alors un certain temps à Göttingen pour travailler avec Hilbert sur des questions de fondements des mathématiques. Après la guerre, Threlfall se retira de 1918 à 1925 (?) au " Gut[8] " de son oncle à Gohlis, près de Dresden, où il était un *Privatgelehrter*[9]. En 1925, il entra à l'école polytechnique de Dresden sous l'influence de W. Treffz, un mathématicien appliqué. Il y passa sa thèse de doctorat et son habilitation ; tous les deux traitent des questions de la théorie des groupes en relation avec des questions topologiques dans le style de Dehn. Threlfall était alors influencé par Friedrich Levi, un mathématicien de Leipzig, auteur d'un traité sur ces questions.[10] Il semble que la connais-

[6] Pour obtenir des surfaces, on travaille dans le plan. Mais Killing a considéré aussi le cas tridimensionnel – sans se rendre compte vraiment des variétés tridimensionnelles fermées (il ne quittait que rarement l'espace tridimensionnel).

[7] On trouvera plus de détails sur les vies de Threlfall et Seifert dans l'étude annoncée (voir note 1).

[8] C'est à dire une terre noble.

[9] Ce terme allemand veut dire " un savant sans but lucratif ".

[10] Dans lequel on trouve entre autre un traitement précis de la classification des surfaces fermées.

sance de Threlfall de la littérature mathématique de son époque était assez limitée.[11] C'est peut-être l'explication du fait que sa démarche était un peu à l'écart des grands courants de la topologie de son époque.

Le premier travail de Threlfall et Seifert[12], paru vers la fin de 1930, commence par une sorte de programme :

" L'énumération de tous les espaces fermées tridimensionnels qui ne sont pas homéomorphes est un problème non résolu. En dimension 2, chaque surface est le domaine de discontinuité [aujourd'hui : domaine fondamental] d'un groupe discret de mouvements qui conservent la metrique soit de la sphère, soit du plan euclidien ou du plan hyperbolique. Ce qui motive l'étude des propriétés topologiques des domaines fondamentaux des groupes discrets tridimensionnels de mouvements de l'hypersphère, de l'espace euclidien ou de l'espace hyperbolique. " [[33], 1]

L'aspect géométrique est présent ici par les groupes de mouvements.

La situation en dimension 2 peut se résumer ainsi : toute surface fermée et orientable est une forme d'espace au sens de Clifford-Klein (précisé par Killing), c'est à dire un espace homogène. Ce résultat présente deux aspects :

- D'abord, il faut déterminer quelles sont les surfaces fermées orientables ?
- Puis il faut démontrer que chacune de ces surfaces est une forme d'espace.

À ma connaissance, on est loin, de nos jours, du résultat envisagé par Threlfall et Seifert même si les cas euclidien (Hantzsche et Wendt [8] après Bieberbach et Nowacki) et sphérique (Threlfall et Seifert) sont bien connus.

Si on désigne par SO(4) le groupe des isométries propres de la sphère S^3, on peut préciser le programme de Threlfall et Seifert comme suit :

- Déterminer tous les sous-groupes de SO(4) dont l'opération sur la sphère est discontinue et sans points fixes.[13]
- Etudier les propriétés topologiques des domaines fondamentaux.
- Donner une classification topologique de ces domaines.

On peut distinguer deux parties dans ce programme de recherche : un problème algébrique (la détermination des sous-groupes finis de $SO(4)$) et un problème géométrique (l'étude des domaines fondamentaux). Pour caractériser les domaines obtenus, on va utiliser de la topologie, en particulier celle du groupe fondamental.

On comprend très bien que le projet de Threlfall et Seifert était assez complexe et que le retour de la géométrie en était substantiel.

[11] Il ne connaît pas même l'article de Hopf sur le problème des formes d'espace (1925), où plusieurs problèmes discutés par Threlfall et Seifert étaient résolus.

[12] La collaboration entre Herbert Seifert (1907-1996), professeur à Heidelberg de 1935 à 1975, et Threlfall fut couronnée par la rédaction de leur traité *Lehrbuch der Topologie* (1934), un manuel qui est encore utilisé aujourd'hui.

[13] Cette hypothèse technique est nécessaire si on veut obtenir la structure d'un revêtement universel. Comme la sphère est compacte, les hypothèses ci-dessus sont équivalentes à l'hypothèse que le groupe considéré soit fini et l'opération libre.

3 Le problème algébrique

Pour déterminer les sous-groupes finis de SO(4), Threlfall et Seifert se servait d'un résultat connu depuis longtemps qui s'exprime en langage moderne par l'affirmation de l'existence d'un morphisme surjectif

$$SO(3) \times SO(3) \to SO(4)$$

Le noyau K de ce morphisme contient deux éléments ; on obtient un isomorphisme (appelé " *the incredible isomorphism* " par B. Thurston) entre $SO(3) \times SO(3)/K$ et $SO(4)$.

Ce résultat était implicite dans le mémoire de E. Goursat " Sur les substitutions orthogonales " ; F. Klein l'a aussi étudié à plusieurs reprises (voir son livre sur l'icosaèdre [[15], chap. I, § 3] et son article sur la géométrie non-euclidienne de 1890). Une discussion détaillée de ce morphisme se trouve dans son livre " Leçons sur la géométrie non-euclidienne " ([16], chap. III, § 3) ; le résultat est cité brièvement dans l'article de Steinitz (1916) sur les polyèdres dans " l'Encyclopédie des Mathématiques ".[14]

Pour établir ce résultat, on peut choisir entre deux possibilités[15] :

- Soit on travaille avec les paramètres dits d'Euler-Cayley et des quaternions en particulier des spineurs. Les paramètres d'Euler définissent un isomorphisme entre $SO(3)$ et S^3 (ou Spin) ; un élément de $SO(4)$ peut se décomposer en deux spineurs (il y a en fait deux manières d'obtenir cette décomposition grâce au signe). D'un point de vue technique, il est peut-être un peu plus commode d'utiliser des matrices unitaires au lieu des matrices orthogonales, c'est à dire des éléments de $SU(2)$. A l'aide des matrices dites de Pauli, il est possible de transformer un spineur en une matrice unitaire en remplaçant les unités quaternioniennes par ces matrices.
- Soit on considère l'espace elliptique et la décomposition d'un mouvement elliptique propre en deux translations de Clifford (une de première espèce, une de seconde espèce) ; les translations de Clifford sont liées à certaines rotations de l'espace ordinaire tridimensionnel. C'est la démarche choisie par Klein dans son livre cité et utilisé plus tard par H. Hopf en 1925.

Threlfall et Seifert ne suivent pas ces deux voies pourtant bien connues à leur époque. En effet, nos deux auteurs ont proposé une troisième solution pour établir l'homomorphisme

$$SO(3) \times SO(3) \to SO(4)$$

fondée sur une décomposition des éléments de $SO(4)$ en un produit de deux facteurs qui, vus du point de vue de la géométrie sphérique, sont une " rotation de droite " et une " rotation de gauche " selon Threlfall et Seifert[16]. Cette décomposition est en fait analogue à celle obtenue avec des translations de Clifford. Dans une base adaptée les rotations de droite

[14] On peut se servir de ce morphisme pour déterminer les polytopes réguliers dans l'espace à quatre dimensions.

[15] [28].

[16] Il semble qu'il n'y a pas de terminologie généralement acceptée dans ce contexte.

s'écrivent sous la forme suivante :

$$\begin{pmatrix} \cos\varphi & -\sin\varphi & 0 & 0 \\ \sin\varphi & \cos\varphi & 0 & 0 \\ 0 & 0 & \cos\varphi & -\sin\varphi \\ 0 & 0 & \sin\varphi & \cos\varphi \end{pmatrix}$$

et les rotations de gauche :

$$\begin{pmatrix} \cos\psi & -\sin\psi & 0 & 0 \\ \sin\psi & \cos\psi & 0 & 0 \\ 0 & 0 & \cos\psi & \sin\psi \\ 0 & 0 & -\sin\psi & \cos\psi \end{pmatrix}$$

Un calcul simple montre qu'on peut toujours décomposer une matrice de $SO(4)$ dans une base adaptée dans un produit d'une rotation de droite avec une rotation de gauche. En général on dispose de deux possibilités qui se distinguent par l'application antipodale (ou – plus simplement – par le signe).

Beaucoup plus compliquée est l'idée avec laquelle Threlfall et Seifert montrent que les rotations de droite[17] et les éléments de $SO(3)$ sont en bijection. Je ne veux pas entrer ici dans les détails qui sont assez compliqués.[18]

Depuis longtemps on connaît les sous-groupes finis de $SO(3)$: ce sont les trois groupes de symétrie des corps platoniciens, les groupes cycliques finis et les groupes diédriques. Mais comment peut-on arriver aux sous-groupes finis du produit $SO(3) \times SO(3)$? Ce problème (le " Gruppenpaarung ") fut résolu par Threlfall et Seifert à l'aide d'un critère qui utilise les sous-groupes distingués des facteurs : Etant donné le produit $G \times H$, soient U et V des sous-groupes de G et de H. S'il existe des sous-groupes distingués u et v de U et de V et un isomorphisme $\alpha : U/u \to V/v$, alors Γ_α - le graphe de α - est un sous-groupe de $U/u \times V/v$.

L'image inverse de ce sous-groupe sous le produit des projections canoniques est un sous-groupe de U × V, donc un sous-groupe de G × H. De plus, on montre que l'on obtient ainsi tous les sous-groupes d'un produit.

Construire tous les sous-groupes finis de SO(3) × SO(3) de cette manière revient à un travail considérable. Mais même les cas les plus simples – on choisit U ={e} et V = v (un sous-groupe fini de SO(3)) – donnent des exemples très intéressants.

Ici nous allons étudier seulement le cas $U = \{e\}$ et $V = \mathbb{Z}/m\mathbb{Z}$; le sous-groupe de $SO(4)$ est alors un groupe cyclique d'ordre $2m$ (si m est impair il y a un autre groupe d'ordre m qui est l'image de $U \times V$).[19]

4 Le problème géométrique : le domaine fondamental

Soit un sous-groupe fini G de $SO(4)$ opérant sur S^3 sans point fixe. Comme l'opération de G sur S^3 est a fortiori discontinue[20] sous cette hypothèse, on peut construire le domaine

[17] Le groupe des rotations de droite sera dans la suite noté SOR et celui des rotation de gauche SOL.

[18] Pour plus de détails, on peut consulter l'étude annoncée en note 1.

[19] Si m est pair le groupe $\mathbb{Z}/m\mathbb{Z}$ contient par un théorème de Cauchy une involution et les formes d'espace elliptiques sont les mêmes que les formes d'espace sphériques.

[20] C'est à dire sans points de cumulation.

fondamental de G. Le procédé par lequel on peut déterminer ce domaine fondamental est connu dès le 19^e siècle (il semble que Dirichlet fut le premier à se servir de cette notion consciemment) : On considère un point O et toutes ses images O_i sous les générateurs de G et sous leurs inverses. On cherche alors les plans de symétrie entre O et O_1, O et O_2 etc. Le domaine fondamental cherché est l'intersection de tous les demi-espaces limités par ces plans de symétrie contenant le point O. Comme nous travaillons dans la sphère S^3, les plans de symétries sont des " grandes-sphères " (c'est à dire des sphères d'unités autour de l'origine, intersection des hyperplans passant par l'origine et de la sphère S^3).

Pour vraiment déterminer le domaine fondamental, il faut connaître l'action du groupe d'un point de vue géométrique.

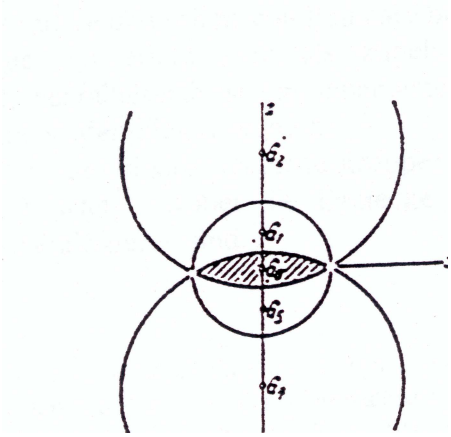
Soit G l'image inverse de $G' = \{e\} \times \mathbb{Z}/m\mathbb{Z}$ (G' est un sous-groupe de $SOL \times SOR$ ou – ce qui revient au même ici – un sous-groupe de SOR). Le générateur de G est une rotation de droite de $SO(4)$ d'ordre $2m$; par conséquent cette isométrie s'écrit comme

$$j = \begin{pmatrix} \cos\frac{\pi}{m} & -\sin\frac{\pi}{m} & 0 & 0 \\ \sin\frac{\pi}{m} & \cos\frac{\pi}{m} & 0 & 0 \\ 0 & 0 & \cos\frac{\pi}{m} & -\sin\frac{\pi}{m} \\ 0 & 0 & \sin\frac{\pi}{m} & \cos\frac{\pi}{m} \end{pmatrix}$$

Autrement dit, il existe une base $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ et des espaces invariants (Eigenräume) $E_{12} = \vec{v}_1\mathbb{R} + \vec{v}_2\mathbb{R}$ et $E_{34} = \vec{v}_3\mathbb{R} + \vec{v}_4\mathbb{R}$ [$\mathbb{R}$ sont les réels], dans laquelle la matrice de l'isométrie est de la forme indiquée. On obtient une interprétation géométrique en projetant stéréométriquement (avec centre (0,0,0,1) dans la base choisie) la sphère S^3 sur un espace de dimension trois représenté par l'hyperplan $\{(x_1, x_2, x_3, x_4) \,|x_4 = 0\}$. Si l'on note x, y, z les coordonnées de l'espace de dimension trois, l'espace invariant $E_{12} = \vec{v}_1\mathbb{R} + \vec{v}_2\mathbb{R}$ est projeté sur le plan d'équation $z = 0$ et le grand-cercle $(E_{12} = \vec{v}_1\mathbb{R} + \vec{v}_2\mathbb{R}) \cap S^3$ se projette sur le cercle d'équation $x^2 + y^2 = 1$. De manière analogue, le grand-cercle $(E_{34} = \vec{v}_3\mathbb{R} + \vec{v}_4\mathbb{R} \cap S^3$ est projeté sur l'axe des z. L'image de l'isométrie est alors un mouvement hélicoïdal autour de l'axe des z, c'est à dire la combinaison d'une rotation autour de l'axe des z et d'une translation parallèle à cette axe (dans la sphère S^3 cette translation est aussi une rotation !). Les images de l'origine – qui est l'image du point $(0, 0, 1, 0)$ sous la projection choisie - se trouvent toutes sur l'axe des z (si m est pair le point à l'infini est aussi une image de l'origine !). Le cercle d'unité $x^2 + y^2 = 1$ est transformé en lui-même (par une rotation de π/m).

L'image décrite est conforme mais ne conserve pas les distances. Par conséquent, il faut retourner à l'hypersphère pour déterminer les plans de symétrie. Pour s'en faire une image concrète, on considère l'axe des z comme " fermé " (on compactifie l'espace tridimensionnel par un point et on obtient une sphère). L'isométrie $\mathcal{J}$ transforme l'axe des z fermé en lui-même ; soit O_1 le point $\mathcal{J}(O)$ $[O = (0, 0, 1, 0)]$. La distance sphérique de O et de O_1 est π/m. Par conséquent le plan de symétrie sphérique passe à la distance $\pi/2$m de O par l'axe des z fermé. Parce qu'un plan sphérique est une grand-sphère (son centre est $(0, 0, 0, 0)$), sa position est bien déterminée. En particulier c'est une sphère qui passe par le grand-cercle projeté dans le cercle $x^2 + y^2 = 1$.

On peut déterminer ainsi tous les plans de symétrie nécessaires ; En projetant dans l'espace tridimensionnel, on obtient :



En général les espaces prismatiques sont les domaines fondamentaux des groupes du type $\{e\} \times G$, où G est un groupe dièdrique binaire[21].

Les plans de symétrie passent tous par le cercle $x^2 + y^2 = 1$. Cette représentation permet de comprendre d'où vient le nom d'espace lenticulaire (lens space ?) ! [22] La variété $L(2m, 1)$ est construite par identification des calottes opposées de la lentille. L'isométrie se décrit comme un mouvement hélicoïdal autour de l'axe des z fermé. L'épaisseur de la lentille est π/m ; l'isométrie se décrit comme la combinaison d'une rotation d'angle π/m de l'axe des z et d'une rotation autour de l'axe des z fermé du même angle[23]. Les points situés sur l'arête de la lentille sont identifiés avec d'autres points de l'arête – l'arête de la lentille est donc un espace invariant.

Le groupe fondamental d'un espace lenticulaire est un groupe cyclique d'ordre $2m$. Il s'ensuit que deux espaces de lentille $L(2m, 1)$ et $L(2m', 1)$ ne sont pas homéomorphes si m n'est pas égal à m'. Si l'on travaille avec deux groupes cycliques on obtient l'espace lenticulaire $L(m, n)$ [39, 40, 41, 42][24].

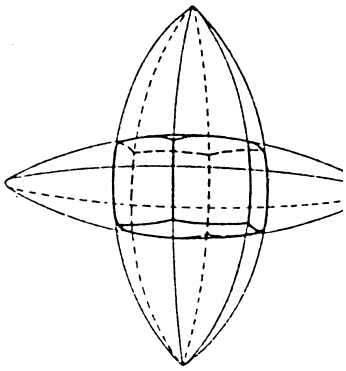
En utilisant les mêmes techniques, Threlfall et Seifert construisirent d'autres variétés tridimensionnelles fermées. Des exemples intéressants sont les espaces prismatiques : ce sont des prismes sphériques admettant deux polygones réguliers comme base et antibase. Les autres faces sont des carrés sphériques. Un cas particulier est l'espace quaternionien déjà étudié par Poincaré ([21], 232-233 et 244-245). Dans le cas de l'espace quaternionien, le prisme est un cube sphérique dont les faces sont identifiées par des mouvements hélicoïdaux d'angle de rotation de 90 ; le groupe qui opère ici est le groupe de Klein binaire (*Kleinsche Vierergruppe*).

[21] " binaire " veut dire que l'ordre du groupe est doublé.

[22] Il est facile de calculer le volume de la variété obtenue : c'est le volume de l'hypersphère (c'est à dire π^2) divisé par l'ordre du groupe G (c'est à dire $2m$).

[23] Pour faciliter la compréhension intuitive, il est important de se convaincre que ces deux rotations sont indépendantes l'une de l'autre.

[24] La classification des espaces lenticulaires fut achevée par Threlfall et Seifert (1932), Reidemeister et Franz (1935) ; [30], 269-273.



Un autre exemple extrêmement important est l'espace dodécaèdral, construit à l'aide du produit $\{e\} \times G$, où G est le groupe dodécaèdral. Ce groupe peut se décrire à l'aide de générateurs et de relations comme suit : $\langle a, b, c | a^5 = b^3 = c^2 = abc \rangle$.

Dans ce cas aussi, on projette sur un espace de dimension 3 de manière que les générateurs soient représentés par les opérations de symétrie (des rotations) du dodécaèdre. Le domaine fondamental de ce groupe est un dodécaèdre sphérique ; le groupe fondamental de la variété est le groupe dit dodécaèdral binaire – un groupe d'ordre 120. Comme ce groupe est identique à son sous-groupe des commutateurs, la variété possède une homologie triviale en dimension un. C'est donc une sphère d'homologie.[25]

Les variétés étudiées par Seifert et Threlfall sont les suivantes[26] :

- Espace projectif (0,0 ;2 ;-)
- Espace de lentille L(p,1) (0,0 ;2p ;-) si p est pair
  (0,0 ;p ;-) et (0,0 ;2p ;-) si p est impair
- Espace prismatique P(m,1) (0,0 ;2,2 ;-) si m est pair
  (0,0 ;4 ;-) si m est impair
- Espace octaédral (0,0 ;2 ;-)
- Cube tronqué (0,0 ;2 ;-)
- Espace de dodécaédal (0 ;0 ;- ;-)

Dans le cas où P n'est pas un point des deux grands cercles qui sont transformés en eux-même, lorsque l'on trace la trajectoire[27] du point P sous l'action de $\mathcal{J}$, $\mathcal{J}^2$ etc. la courbe qui en résulte sera fermée parce que l'on a $\mathcal{J}^{2m} = id$. Cette courbe fermée fait un tour complet autour des deux axes. Elle se trouve sur un tore – par conséquent d'un point de vue topologique c'est un nœud toroïdal noté de nos jours $(1, 1)$. On trouve ici l'origine de la

[25] Pour l'histoire de cet exemple important on peut consulter [30], 158–177 ; des détails mathématiques se trouvent chez Weber/Seifert 1932.

[26] (x,y ;u,v) signifient : nombre de Betti en dimension 1, dito en dimension 2 ; coefficients de torsion en dimension 1, dito en dimension 2. On a x = y par dualité et v = 0 parce que les variétés sont orientables.

[27] Pour y arriver on introduit le paramètre t dans la matrice ; autrement dit, on travaille avec l'angle $t\phi$ où t varie entre 0 et 1.

théorie des variétés dites fibrées (ou feuilletées) selon Seifert.[28]

5 La théorie de Seifert

Comment décrire la construction d'un espace lenticulaire ? Les deux axes sont des cercles ordinaires.[29] Autour de ces deux axes, il y a des tores concentriques. Ces tores sont fibrés par des nœuds de même type (par exemple du type (1,1)). Comme le complément d'un tore solide de S^3 est lui-même un tore solide[30], l'espace lenticulaire est composé de deux tores. Pour bien comprendre l'intersection des deux tores, il faut remarquer que le nœud (m,n) autour d'un axe est le noeud (n,m) autour de l'autre.

Herbert Seifert a étudié les variétés tridimensionnelles fermées et orientables munies d'une telle structure dans sa deuxième dissertation présentée à Leipzig sous la direction de B. L. van der Waerden en 1932.[31] L'idée centrale de ce mémoire est la suivante : pour étudier les variétés admettant une structure analogue à celle de l'espace lenticulaire, il faut chercher des invariants qui caractérisent les composants de cet espace, c'est-à-dire des tores solides. Il existe des situations dans lesquelles ces invariants sont si puissants qu'ils forment non seulement des conditions nécessaires mais aussi des conditions suffisantes pour l'homéomorphisme.

Sans entrer dans les détails de cette théorie, je vais seulement indiquer les idées principales. Comme nous l'avons vu précédemment, on peut décrire la structure des formes d'espace sphériques par le fait qu'elles sont composées des tores. Ces tores portent une structure fine : à une exception près, les points du tore se trouvent sur des trajectoires qui sont des nœuds toroïdaux. Les points qui ne vérifient pas cette propriété sont ceux situés sur le cercle au " milieu " du tore (en allemand, l'âme du tore). Seifert décrit ce type de structure sous le terme de " tore fibré" : Soit un cylindre solide dans l'espace ordinaire avec des segments comme indiqués ci-dessous

Si on identifie la base avec le cercle situé au sommet par une rotation de $\frac{2m}{n}\pi$ (m et n des entiers sans diviseur commun, $n \neq 0$), on obtient un tore fibré. Les segments ainsi transformés en nœuds du type (n, m) sont les fibres ordinaires. Le segment au milieu du cylindre, transformé en un cercle sans nœud, est la fibre exceptionnelle. Les nombres m et n sont les nombres caractéristiques du tore fibré.

Une variété est fibrée (selon Seifert) si elle est localement un tore tordu.[32] Autrement dit : les points de la variété sont distribués sur des noeuds bien composés. Par conséquent la variété admet une opération de S^1. Si la variété est compacte, il n'y a qu'un nombre fini de

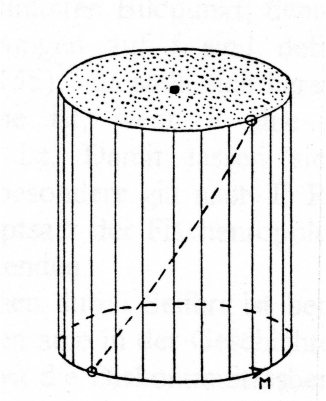
[28] On trouve déjà cette idée (au moins le principe de celle-ci) dans les " Leçons " de Félix Klein. Du point de vue de la géométrie elliptique, la trajectoire d'un point (non situé sur les axes) est une courbe fermée contenue dans un hyperboloïde à une nappe [[16], 125-127].

[29] C'est à dire des nœuds triviaux.

[30] Cette idée fondamentale se trouve pour la première fois chez Dyck en 1884.

[31] La première dissertation fut présentée à Dresden sous la direction de Kowalewski en 1931 sous le titre " Konstruktion dreidimensionaler geschlossener Räume ". On y trouve le théorème dit aujourd'hui de Seifert et van Kampen. A Dresden, Seifert a obtenu le titre de docteur es sciences techniques ; à Leipzig, il reçut celui de docteur es philosophie (considéré supérieur dans les cercles académiques en Allemagne).

[32] Si on veut être précis ici il faut préciser la notion de l'équivalence fibre par fibre, c'est-à-dire il faut faire attention à la structure fine de la variété.



fibres exceptionnelles (ce qui signifie que la variété est composée d'un nombre fini de tores fibrés).

L'idée la plus novatrice de Seifert est peut-être celle de la surface de découpage (" Zerlegungsfläche ") : chaque fibre de la variété est représentée sur la surface de découpage par un point ; la topologie de la surface de découpage est la topologie aujourd'hui appelée finale sous la projection canonique. De nos jours, la surface de découpage est introduite comme un espace quotient – une idée assez hardie en 1932 ! La surface de découpage fut peut-être le premier espace topologique défini de manière tout à fait abstraite.

Seifert démontre que si la variété fibrée est fermée, sa surface de découpage est fermée aussi. La surface de découpage contient des informations sur la structuration de la variété par les fibres. Seifert démontre que deux variétés fibrées de la même classe – ce qui signifie qu'elles sont homéomorphes fibres par fibres – ont toujours des surfaces de découpage homéomorphes. Par conséquent, l'indice d'Euler-Poincaré de la surface de découpage est un invariant des variétés tridimensionnelles.

Ainsi, Seifert obtient les invariants suivants d'une variété tridimensionnelle fermée fibrée :

$$(O/N; o/n; g|b; n_1,_1\,; n_2, m_2; \ldots; n_k, m_k)$$

où O/N symbolise l'énoncé " la variété est orientable [O] ou non [N] ", o/n l'énoncé " sa surface de découpage est orientable [o] ou non [n] ", b est un entier défini par une sorte de chirurgie ; g est l'indice d'Euler-Poincaré de la surface de découpage et n_i,m_i sont les nombres caractéristiques des fibres exceptionnelles ($i = 1, \ldots, k$).

Il y a des situations où ces invariants sont suffisants pour classer les variétés fibrées même du point topologique : ce sont les variétés fibrées dont la surface de découpage est homéomorphe à une sphère. En particulier, Seifert peut démontrer que la conjecture dite de Poincaré[33] est vraie pour les variétés fermées tridimensionnelles fibrées : une telle variété simplement connexe est homéomorphe à la sphère S^3.

Pour obtenir ce résultat, Seifert démontre qu'une variété fibrée dont la surface de découpage est homéomorphe à une sphère, possède un groupe fondamental fini si et seulement si le nombre de fibres exceptionnelles est égal à 3 ou moins. Or les variétés fibrées avec un tel

[33] [30], 239 - 269.

nombre de fibres exceptionnelles peuvent se construire par la méthode de Threlfall et Seifert. Elles sont donc des formes d'espace sphérique analysées dans le travail de 1930. Par conséquent, il suffit d'étudier ces formes et on trouve que parmi celles-ci la seule forme simplement connexe est la sphère S^3.

6 Quelques remarques comme conclusion

On peut certes donner un exposé plus moderne de la théorie de Seifert et des travaux de Threlfall et Seifert beaucoup plus précis et plus court (on trouve par exemple un tel exposé dans [3]). Mais du point de vue de l'historiographie, ce n'est pas le but principal. Si on veut savoir comment les mathématiques se développent il est important aussi d'en suivre les détours et même les dédales.

Le plus remarquable à mon avis dans cette histoire, c'est l'intégration de plusieurs aspects dans une seule théorie : nous avons rencontré la géométrie classique (sphérique et elliptique), la théorie combinatoire des groupes (générateurs et relations), la topologie générale (espace homogène, revêtement, surface de découpage) et la topologie algébrique (groupe fondamental) ... La réintégration de la géométrie dans le cadre de la théorie des variétés est un fait remarquable, parce que la période antérieure du développement de la topologie était caractérisée par un processus qu'on peut nommer peut-être " décortiquer la topologie ". On essayait de démontrer le plus possible sans avoir recours à la géométrie, à l'analyse etc. (c'était un procédé de " décontextualisation "). Seulement après avoir ainsi déterminé les limites de la topologie on était prêt au retour de la géométrie !

Bibliographie

[1] L. Bieberbach, *Über die Bewegungen der euklidischen Räume I* , Mathematische Annalen, **70**, 1911, 297-336.

[2] L. Bieberbach, *Über die Bewegungen der euklidischen Räume II*, Mathematische Annalen, **72**, 1912, 400-412.

[3] CMc. A. Gordon, *Three-dimensional Topology up to 1960*, In *History of Topology*, ed. by I. M. James, Amsterdam : Elsevier, 1999, 449-489.

[4] M. Dehn and P. Heegrad, *Analysis situs*, In *Enzyklopädie der mathematischen Wissenschaften mit Einschluß ihrer Anwendungen*, Dritter Band in drei Teilen : Geometrie. Redigiert von W. Fr. Meyer und H. Mohrmann. Erster Teil, erste Hälfte Leipzig : Teubner, 1907-1910, 153-220, abgeschlossen 25-06-1907.

[5] P. G. Lejeune Dirichlet, *Ueber die Reduction der positiven quadratischen Formen mit drei unbestimmten Zahlen*, Journal für die reine und angewandte Mathematik, **40**, 1850, 209-227.

[6] W. Dyck, *On the, Analysis situs" of Three-dimensinal Spaces*, In : *Report of the fifty-forth Meeting of the British Association for the Advancement of Science*, held at Montreal in August and September 1884 London, 1885, 648.

[7] W. Franz, *Über die Torsion einer Überdeckung*, Journal für die reine und angewandte Mathematik, **173**, 1935, 245-254.

[8] W. Hantzsche and H. Wendt, *Dreidimensionale euklidische Raumformen*, Mathematische Annalen, **110**, 1935, 593-611.

[9] W. Hantzsche and H. Wendt, *Dreidimensionale wirbelgefaserte Räume*, Mathematische Zeitschrift, **40**, 1936, 727-755.

[10] J. Hempel, *3-manifolds*, Princeton [N.Y.] : Princeton University Press, 1976.

[11] *History of Topology*, ed. by I. M. James, Amsterdam : Elsevier, 1999.

[12] H. Hopf, *Das Klein-Cliffordsche Raumproblem*, Mathematische Annalen, **95**, 1926, 313-339.

[13] I. M. James, *From combinatorial topology to algebraic topology*, In *History of Topology*, ed. by I. M. James, Amsterdam : Elsevier, 1999, 561-574.

[14] F. Klein, *Ueber nichteuklidische Geometrie*, Mathematische Annalen, **37**, 1890, 544 – 572.

[15] F. Klein, *Vorlesungen über das Ikosaeder und die Auflösung der Gleichungen fünften Grades. Herausgegeben mit einer Einführung und Kommentaren versehen von P. Slodowy*, Basel u.a. : Birkhäuser, 1993 ; Erstausgabe Leipzig : Teubner, 1884.

[16] F. Klein, *Vorlesungen über nicht-euklidische Geometrie*, Berlin : Springer, 1928 ; unveränderter Nachdruck Berlin u.a. : Springer, 1968.

[17] H. Kneser, *Geschlossene Flächen in dreidimensionalen Mannigfaltigkeiten*, Jahresbericht der Deutschen Mathematiker Vereinigung, **38**, 1929, 248 – 260.

[18] Fr. Levi, *Geometrische Konfigurationen. Mit einer Einführung in die kombinatorische Flächentopologie*, Leipzig, 1929.

[19] W. Nowacki, *Die euklidischen, dreidimensionalen, geschlossenen und offenen Raumformen*, Commentarii mathematicae Helveticii, **7**, 1934, 81-92.

[20] P. Orlik, *Seifert-manifolds*, Berlin u.a. : Springer, 1972.

[21] H. Poincaré, *Oeuvres*, Tome **6**, Paris : Gauthier-Villars, 1953.

[22] D. Puppe, *Herbert Seifert Nachruf In : Jahrbuch der Heidelberger Akademie der Wissenschaften für 1997*, Heidelberg u.a. : Springer, 1998, 139-144.

[23] D. Puppe, *Herbert Seifert*, In *History of Topology*, ed. by I. M. James, Amsterdam : Elsevier, 1999, 1021-1027.

[24] J. G. Ratcliffe, *Foundations of Hyperbolic Manifolds*, New York u.a. : Springer, 1994.

[25] K. Reidemeister, *Überdeckungen von Komplexen*, Journal für die reine und angewandte Mathematik, **173**, 1935, 164-173.

[26] S. L. Segal, *Topologists in Hitler's Germany*, In *History of Topology*, ed. by I. M James, Amsterdam : Elsevier, 1999, 849-861.

[27] E. Steinitz, *Polyeder und Raumteilungen*, In *Enzyklopädie der Mathematischen Wissenschaften mit Einschluß ihrer Anwendungen.* Dritter Band : Geometrie, redigiert von W. Fr. Meyer und H. Mohrmann. Erster Teil, zweite Hälfte Leipzig, 1914-1931, 1-139, ausgegeben 31-8-1916.

[28] C. Tzanakis, *Rotations, nombres complexes et quaternions*, L'Ouvert No. **75**, juin 1994, 15 – 31.

[29] B. L. van der Waerden, *Kombinatorische Topologie*, Jahresbericht der Deutschen Mathematiker-Vereinigung, **39**, 1930, 121-139.

[30] K. Volkert, *Das Homöomorphierproblem insbesondere der 3-Mannigfaltigkeiten in der Topologie 1895-1935*, Paris : Kimé, 2002.

[31] K. Volkert, *Wie und warum wurde Poincaré zum Topologen ?*, Philosophia scientiae, **2**, 1997, 74-102.

[32] W. Threlfall, *Bedingt konvergente Reihen*, Mathematische Zeitschrift, **24**, 1926, 212-214.

[33] W. Threlfall and H. Seifert, *Topologische Untersuchung der Diskontinuitätsbereiche endlicher Bewegungsgruppen des dreidimensionalen sphärischen Raumes*, Mathematische Annalen, **104**, 1931, 1 - 70.

[34] H. Seifert, *Konstruktion dreidimensionaler geschlossener Räume*, Verhandlungen der Sächsischen Akademie der Wissenschaften zu Leipzig. Mathematisch-physikalische Klasse, **83**, 1931, 26-66.

[35] H. Seifert, *Topologie dreidimensionaler gefaserter Räume*, Acta mathematica, **60**, 1932, 147-238.

[36] H. Seifert, *Homologiegruppen berandeter dreidimensionaler Mannigfaltigkeiten*, Mathematische Zeitschrift, **35**, 1932, 609-611.

[37] H. Seifert, *Poincaré'sche Räume*, In *Verhandlungen des Internationalen Mathematiker-Kongresses Zürich 1932*. II. Band : Sektionsvorträge, hg. von W. Saxer Zürich und Leipzig o. J., 197-198.

[38] H. Seifert, *Lösung der Aufgabe 84*, Jahresbericht der Deutschen Mathematiker-Vereinigung, **41**, 1932, 7.

[39] W. Threlfall, *Dreidimensionale Raumformen*, In *Verhandlungen des Internationalen Mathematiker-Kongresses Zürich 1932*. II. Band : Sektionsvorträge, hg. von W. Saxer Zürich und Leipzig o. J., 148-149.

[40] W. Threlfall, *Räume aus Linienelementen*, Jahresbericht der Deutschen Mathematiker-Vereinigung, **42**, 1932, 87-110.

[41] W. Threlfall, *Lösung der Aufgabe 84*, Jahresbericht der Deutschen Mathematiker-Vereinigung, **41**, 1932, 6-7.

[42] W. Threlfall and H. Seifert, *Topologische Untersuchung der Diskontinuitätsbereiche endlicher Bewegungsgruppen des dreidimensionalen sphärischen Raumes*, Schluß, Mathematische Annalen, **107**, 1932, 543-586.

[43] H. Seifert, *Verschlingungsinvarianten*, Preußische Akademie der Wissenschaften. Sitzungsberichte der physikalisch-mathematischen Klasse, 1933, 811-823.

[44] H. Seifert, *Aufgaben Nr. 154, 236*, Jahresbericht der Deutschen Mathematiker-Vereinigung, verschiedene Bände.

[45] H. Seifert, *Lösung der Aufgabe 124*, Jahresbericht der Deutschen Mathematiker-Vereinigung, **42**, 1933, 112-113.

[46] W. Threlfall, *Aufgaben Nr. 84, 135, 189*, Jahresbericht der Deutschen Mathematiker-Vereinigung, verschiedene Bände.

[47] W. Threlfall, *Lösung der Aufgabe 124*, Jahresbericht der Deutschen Mathematiker-Vereinigung, **42**, 1933, 116-117.

[48] C. Weber and H. Seifert, *Die beiden Dodekaederräume*, Mathematische Zeitschrift, **37**, 1933, 237 – 253.

[49] H. Seifert, *Über das Geschlecht von Knoten*, Mathematische Annalen, **110**, 1934, 571-592.

[50] H. Seifert and H. Threlfall, *Lehrbuch der Topologie*, Leipzig : Teubner, 1934. Englische Ausgabe : Seifert and Threlfall : A Textbook of Topology and Seifert : Topology of 3-dimensional fibered spaces, ed. by J. Birman and J. Eisner, New York u.a. : Academic Press, 1980. Weitere Ausgaben in Spanisch, Russisch und Chinesisch.

[51] W. Threlfall, *La notion de recouvrement*, L'Enseignement mathématique, **34**, 1935, 228-254.

[52] H. Seifert, *Die Verschlingungsinvarianten der zyklischen Knotenüberlagerungen*, Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität, **11** 1936, 84 – 101.

[53] H. Seifert, *Algebraische Approximation von Mannigfaltigkeiten*, Mathematische Zeitschrift, **41**, 1936, 1-17.

[54] H. Seifert, *La théorie des noeuds*, l'Enseignement mathématique, **35**, 1936, 201-212.

[55] H. Seifert, *Bemerkung zur stetigen Approximation von Flächen*, Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität, **12**, 1936, 29-37.

[56] W. Threlfall, *Quelques résultats récents de la topologie des variétés*, L'Enseignement mathématique 35 1936, 242-255 & notes complémentaires.

[57] W. Threlfall, *Modern mathematics in Germany*, The Mathematics Teacher, **29**, 1936, 300-309.

[58] H. Seifert and W. Threlfall, *Variationsrechnung im Großen*, Morsesche Theorie [Hamburgische Mathematische Einzelschriften **24**, Heft/1938] Leipzig und Berlin : Teubner, 1938.

[59] W. Threlfall, *Le calcul des variations globales*, L'Enseignement mathématique, **38** 1939-42, 189-208.

[60] W. Threlfall, *Stationäre Punkte auf geschlossenen Mannigfaltigkeiten*, Jahresbericht der Deutschen Mathematiker-Vereinigung, **51** 1941, 14-33.

[61] H. Seifert, **Lösung der Aufgabe 307**, Jahresbericht der Deutschen Mathematiker-Vereinigung, **52**, 1942, 72-79.

[62] H. Seifert, *Zur assymptotischen Integration von Differentialgleichungen*, Mathematische Zeitschrift, **48**, 1942, 175-192.

[63] H. Seifert, *Die hypergeometrischen Differentialgleichungen der Gasdynamik*, Mathematische Annalen, **120** 1947, 75-126. [Festschrift für C. Caratheordory]

[64] H. Seifert, *Periodische Bewegungen mechanischer Systeme*, Mathematische Zeitschrift, **51**, 1948, 197-216.

[65] H. Seifert and W. Threlfall, *Topologie. In : Naturforschung und Medizin in Deutschland 1939-1945*, Für Deutschland bestimmte Ausgabe der FIAT Reviews of German Science. Band **2** : Reine Mathematik. Teil 2, hg. von W. Süss, Wiesbaden : Steiner, 1948, 239-252. Englisches Original 1946.

[66] W. Threlfall, *Knotengruppe der Homologieinvarianten* Sitzungsberichte der Heidelberger Akademie der Wissenschaften, Mathematisch-naturwissenschaftliche Klasse, 1949, **8**, Abhandlung, 357-370.

[67] H. Seifert, *Schlingknoten Mathematische Zeitschrift*, **52**, 1950, 62-80.

[68] H. Seifert, *Closed integral curves in 3-space and isotopic two dimensional deformations*, Proceedings of the American Mathematical Society **1** 1950, 287-302.

[69] H. Seifert, *On the homology invariants of knots*, Quaterly Journal of Mathematics. Oxford Series, **2**, Serie 1, 1950, 23-32.

[70] H. Seifert and W. Threlfall, *Old and new results on knots*, Canadian Journal of Mathematics **2**, 1950, 1-15.

[71] H. Seifert, *Zum Satz von O. Bonnet über den Durchmesser einer Eifläche*, Mathematische Zeitschrift **77**, 1961, 125-130.

[72] H. Seifert, *Minimalflächen von vorgegebener topologischer Gestalt* Sitzungsberichte der Heidelberger Akademie der Wissenschaften, Mathematisch-naturwissenschaftliche Klasse Jahrgang 1974, 1. Abteilung, **16** Seiten.

[73] H. Seifert, *Komplexe mit Seitenzuordnung Nachrichten der Akademie der Wissenschaften in Göttingen. II.*, Mathematisch-physikalische Klasse Jahrgang 1975, No. **6**, 49-80.

[74] V. Klaus, Institut für Didaktik der Mathematik, Fachbereich Mathematik, Johann Wolfgang Goethe-Universität Frankfurt a.M., Senckenberganlage **9**, D-60054 Frankfurt a.M.

Histoire des immersions isométriques

Abdelghani Zeghib[1]
CNRS, UMPA, École Normale Supérieure de Lyon
46, allée d'Italie, 69364 Lyon cedex 07, FRANCE

1 Introduction

Ceci est un essai pour retracer l'évolution historique d'un problème central et naturel en géométrie, celui des immersions (et plongements) isométriques. L'auteur avoue et insiste sur son incompétence du point de vue chronologique et culturel (i.e. qu'il n'a point entrepris une recherche documentaire systématique). Il pense cependant que ce travail reste intéressant à accomplir, par des gens plus compétents en la matière, i.e. des historiens des mathématiques.

Je voudrais ici en profiter pour exprimer un avis, partagé par certains collègues, sur cette question : un mathématicien (qui n'est pas historien des mathématiques) est-il habilité à parler de l'histoire de son thème ? Mon approche est, sans croire en un déterminisme (strict) dans l'évolution de la recherche mathématique, qu'il est intéressant de permettre à un mathématicien pratiquant, ayant quelques repères historiques, d'interpoler, et de reconstituer une histoire idéale du thème de recherche qui est le sien. Cela donnerait l'histoire non pas telle qu'elle s'est déroulée, mais "comme elle aurait dû logiquement se dérouler" [2]. L'évolution de la recherche d'un thème mathématique est certainement monotone (par définition, disait R. Thom), mais il me paraît qu'il existe, généralement, une sorte de "self-similarité", qui fait que chaque chercheur, retrouve personnellement, au fur et à mesure, les idées qui se sont dégagées au cours de l'histoire, dans son thème. Dans ce cas, le chercheur n'est pas seulement un bon modélisateur, mais aussi un témoin naturel, au service de l'historien, du rôle et de l'importance de différentes idées apparues dans l'évolution de son thème.

À l'opposé des textes historiques standards, le présent texte traite le sujet jusqu'aux développements les plus récents. Ce texte est conçu essentiellement pour exposer les diverses contributions de M. Gromov dans ce domaine. M. Gromov s'est intéressé aux immersions

[1] Zeghib@umpa.ens-lyon.fr
http ://umpa.ens-lyon.fr/~zeghib/

[2] après tout, le mathématicien modélise des phénomènes de toute nature, en tant que systèmes dynamiques, pour prévoir leur futur, et connaître leur passé.

isométriques dès le début de sa carrière (voir son doctorat publié fragmenté en plusieurs articles autour de 1970). Il y introduit, en particulier sa théorie du h-principe permettant de montrer l'abondance des immersions isométriques (aussi bien dans le sens sous-entendu ici, i.e. dans le cas riemannnien, que dans le cas de structures géométriques plus générales). J'ai eu la chance de suivre un cours assuré par Gromov sur ce thème, en 1981-82 (à Paris). En 1986, Gromov publie le livre [14], duquel on puisera beaucoup de faits mathématiques cités ici, au lieu de se référer aux publications originales, car ce livre offre une approche globale et simplifiée.

2 Début de la géométrie extrinsèque

Carl Friedrich Gauss (1777-1855) inaugure au début du 19 ème siècle, l'étude de la géométrie différentielle et métrique (locale) des surfaces dans l'espace euclidien (3-dimensionel). Il a ainsi introduit la première forme fondamentale (i.e. la métrique induite), quantité, à l'évidence extrinsèque, mais il s'est en même temps posé le problème de déformation rigide, ou plus généralement, d'existence d'applications isométriques, entre surfaces, qui soient non-triviales, i.e. qui ne proviennent pas d'un déplacement de l'espace ambiant. Il a aussi introduit, la seconde forme fondamentale, une entité également extrinsèque, mais donnant lieu, à sa grande surprise, à un invariant intrinsèque, qui est la courbure (Theorema Egregium). Il reconnaissait, en particulier, une surface à courbure égale à -1, la pseudo-sphère. Il s'agit de la surface graphe de la fonction

$$(x, y) \to -\sqrt{1 - x^2 - y^2} + \ln \frac{1 + \sqrt{1 - x^2 - y^2}}{\sqrt{x^2 + y^2}}$$

3 Naissance de la géométrie non-euclidienne

En parallèle avec les travaux de Gauss en géométrie métrique des surfaces, on n'a pas cessé de travailler sur la géométrie euclidienne plane synthétique, notamment du point de vue de l'indépendance des postulats d'Euclide , plus précisément, la dépendance du cinquième postulat de ces précédents. Il s'agit bien sûr de celui stipulant que, d'un point, on peut mener une unique droite parallèle à une droite donnée.

Les arguments pour démontrer ce postulat procédaient par l'absurde, pour arriver à un fait dont l'effet choquant est plus évident. En d'autres termes, on développait et étudiait une géométrie hypothétique (imaginaire), dont on espérait montrer un jour l'évidence de son inexistence.

On avait ainsi développé la trigonométrie de cette géométrie et Gauss avait remarqué que cette trigonométrie (pour des petits triangles) coïncidait avec celle de la pseudo-sphère. Il était seulement gêné par la topologie (non-simplement connexe) et l'incomplétude de cette pseudo-sphère, pour en arriver au but inattendu, i.e. la réalité de la géométrie "imaginaire".

Ce n'est qu'autour de 1830, que fût déclarée l'existence de cette géométrie imaginaire, par Jànos Bolyai (1802-1860, fils du plus célèbre Wolfang Bolyai) et Ivanovich Lobatchevsky (1792-1856), et ensuite l'aval par Gauss de cette découverte. Cette existence n'a pas été démontré de façon formelle (par exemple, en donnant un modèle) mais, en exprimant (avec

un courage de braves) que la guerre millénaire contre cette géométrie "imaginaire" était injuste et vaine. Il faudrait au contraire investir ses propriétés, et poursuivre sa construction "positive" tout comme celle de la géométrie euclidienne, et non plus dans le but de la faire s'écrouler (au profit de cette dernière).

4 Géométrie riemannienne

En introduisant les métriques (dites par la suite) riemanniennes, Riemann révolutionne la géométrie, d'une part, en abandonnant tous les postulats d'Euclide (ce qui revient en termes modernes, à essentiellement relaxer l'hypothèse d'homogéinéité locale d'une géométrie ; la contre-révolution à cette idée fût entreprise, dans un certain sens par F. Klein dans son programme d'Erlangen) ; et d'autre part, en munissant une surface (en se restreignant au cas 2-dimensionel) d'un champ de formes quadratiques définies positives, qui ne correspond pas nécessairement à la première forme fondamentale d'une réalisation de cette surface dans l'espace.

5 Modèles de géométrie hyperbolique

Il a donc fallu attendre l'introduction des métriques riemanniennes pour montrer l'existence de la géométrie non-euclidienne plane, et ce en l'exprimant comme métrique riemannienne sur le disque. Ceci fût réalisé par E. Beltrami en 1868, donc presque 40 ans après la "naissance" de cette géométrie. Ceci donne donc une preuve formelle du fait que cette nouvelle géométrie est aussi consistante (ou inconsistante) que la géométrie euclidienne elle même. F. Klein introduit ensuite son modèle projectif (1871), et revenant à l'approche synthétique. H. Poincaré (1854-1912), introduit son modèle conforme, de la géométrie non-euclidienne, en tant que métrique riemannienne sur le disque. Mais le modèle le plus simple dû également à Poincaré , et dit (maintenant) demi-plan de Poincaré noté $\mathfrak{H}^2$, est celui où la métrique est définie sur le demi-plan supérieur de $\mathfrak{R}^2$ par le formule :

$$g = \frac{dx^2 + dy^2}{y^2}$$

Désormais, on utilisera la terminologie moderne *géométrie hyperbolique* (plane) pour désigner cette géométrie non-euclidienne, même si l'auteur ignore l'historique de cette transformation terminologique (à propos de l'évolution terminologique, il est remarquable, qu'un siècle après, M. Gromov introduit une notion de groupes (et espaces) hyperboliques, au sens large, au point où le terme "variété hyperbolique" devient un peu confus (tout comme le terme "géométrie non-euclidienne" était un peu péjoratif!).

6 Problème d'immersions isométriques

On a ainsi vérifié que la pseudo-sphère est localement isométrique au demi-plan de Poincaré . Il paraît ainsi naturel de se poser la question :

"Y-a-t-il une surface dans l'espace (euclidien de dimension 3) qui soit globalement isométrique au demi-plan de Poincaré ?"

De surcroît, la question peut se poser pour une surface quelconque munie d'une métrique riemannienne (au lieu du demi-plan de Poincaré). Une réponse positive à cette question reviendrait à dire que Riemann n'a pas inventé plus de surfaces que Gauss .

7 Obstructions : Théorème de Hilbert

Heureusement (ou malheureusement) Hilbert (1901) [18] a montré que la réponse à cette question est négative. Plus précisément, il n'existe pas dans l'espace une surface, de classe C^2, qui soit globalement isométrique au demi-plan de Poincaré. Faisons, un saut dans le temps pour citer quelques développements de la théorie des obstructions aux immersions isométriques. D'abord l'énoncé de Hilbert a été généralisé, par Efimov (1964) [6] (et revu par T. Milnor (1972) [25], voir aussi Schlencker [32] pour une généralisation dans un autre sens) en remplaçant, le demi-plan de Poincaré par une surface complète à courbure pincée entre deux nombres (strictement) négatifs

Cependant, si *le demi-plan de Poincaré peut se plonger isométriquement dans l'espace euclidien* $\mathfrak{R}^4$. En revanche, le théorème d'Efimov est faux dans ce cas là.

D'après Rosendorn (1960) $\mathfrak{H}^2$ *s'immerge* isométriquement dans $\mathfrak{R}^5$. Plus généralement, l'espace hyperbolique $\mathfrak{H}^n$ s'immerge isométriquement dans $\mathfrak{R}^{4n-3}$.

Blanusa (1955) a démontré que $\mathfrak{H}^2$ se plonge isométriquement de façon C^∞ dans $\mathfrak{R}^6$, et ce comme graphe d'une application $\mathfrak{R}^2 \to \mathfrak{R}^4$. M. Gromov démontre (sous forme d'exercise ([14], p 277) qu'il existe un plongement isométrique analytique (et propre) de $\mathfrak{H}^2$ dans $\mathfrak{R}^8$.

On voit ainsi que les questions autour du théorème de Hilbert, qui peuvent être vues comme la motivation originelle de la théorie (des immersions isométriques) ne sont pas encore réglées.

8 Courbure positive, cas des surfaces : rigidité et existence

Les travaux sur la géométrie extrinsèque des surfaces ne se sont pas arrêtés avec Gauss . Citons, à titre d'exemple, l'œuvre colossale de G. Darboux sur ce sujet, au $19^{\text{ème}}$ siècle, et qui reste source d'inspiration jusqu'à nos jours.

L'une des questions posées depuis cette époque était de voir si certaines surfaces sont rigides, ou au contraire, admettent des deformations isométriques non-triviales (i.e. qui ne soient pas déduites de rotations de l'espace ambient). Ainsi, Cohn-Vossen [4] (voir aussi [3]) démontre la rigidité de la sphère canonique dans l'espace euclidien. Il démontre en fait la rigidité globale de celle-ci (i.e. pas seulement par déformation), en montrant précisément que toute surface complète (ou disons compacte) à courbure positive constante dans l'espace euclidien, est une sphère "ronde". La raison essentielle de cette propriété est la convexité (globale) des surfaces compactes à courbure positive (même variable). Le résultat optimal de rigidité des hypersurfaces de dimension quelconque est attribué à R. Sacksteder (1962) [31].

Dans le cas de surfaces à courbure positive, on doit à Alexandrov (1944), le commencement de la preuve du résultat d'existence et unicité suivant, avec des contributions de Nirenberg (1953) et Pogorelov. Ce dernier en donne la version définitive en (1969) (où l'on ne suppose pas la métrique partout à courbure positive) :

Si une métrique C^∞ (resp. analytique) sur la sphère S^2 admet une courbure $> k$, alors, elle se plonge de façon C^∞ (resp. analytique) dans l'espace 3-dimensionnel universel (i.e. l'espace complet et simplement connexe) à courbure constante k. Ce plongement est rigide (i.e. unique à congruence près) (voir [14], p 289).

9 Obstructions. cas générique

Le problème de l'immersion isométrique d'une variété riemannienne (V^n, g) dans un espace euclidien $\mathfrak{R}^q$ se traduit localement, par une application $(x_1, \ldots, x_n) \to (f_1, \ldots, f_q)$, vérifiant le système de $n(n+1)/2$ équations à q inconnues :

$$\langle \frac{\partial f_k}{\partial x_i}, \frac{\partial f_k}{\partial x_j} \rangle = g_{ij}$$

Ainsi pour une métrique g générique, il n'existe pas de telle immersion s'il y a plus d'équations que d'inconnues, i.e. $q < n(n+1)/2$.

10 Cas analytique : Théorème de Cartan-Janet, systèmes différentiels, et autres approches

Janet [21] démontre en 1926 que, dans le cas analytique (i.e. V et g sont analytiques réelles), le système d'équations des immersions isométriques admet des solutions **locales** dès que $q \geq n(n+1)/2$.

Ce ne serait pas une "erreur historique" d'affirmer qu'Elie Cartan a résolu tous les problèmes **locaux** de la géométrie, à la seule réserve près, d'insister sur le caractère analytique de ses problèmes. En effet, il a développé pour cela sa théorie des systèmes différentielles extérieurs, et a en particulier démontré le théorème d'intégrabilité dit "Théorème de Cartan-Kähler". Cette théorie reste jusqu'à nos jours difficile à pénétrer ! Son contenu analytique est simplement le Théorème de Cauchy-Kowalewka classique, mais ses subtilités algébriques sont "transcendantes".

E. Cartan a effectué les vérifications algébriques, permettant d'appliquer le théorème de Cartan-Kähler, au problème d'immersions isométriques, et a retrouvé ainsi le résultat de Janet (1927). Dans l'article en question [2], Cartan affirme qu'il disposait de cette preuve depuis plusieurs années (on peut en tirer qu'il sous-entendait, avant la parution de celle de Janet). Il rappelle également que c'est Schläffli, qui a énoncé ce théorème pour la première fois en 1871.

Ceci n'a pas empêché la parution de beaucoup d'autres publications plus modernes, simplifiant le travail algébrique d'E. Cartan (voir par exemple [10]).

H. Jacobowitz est parmi ceux ayant fait une brillante carrière d'isométrico-plongeur, tant dans le cas analytique que le cas lisse (voir ci-dessous les répercussions inattendues

du travail de Nash). Dans le cas analytique [19, 20], il applique l'esprit de la preuve du théorème de Cartan-Khäler au problème d'immersions isométriques (ainsi que des manipulations algébriques et différentielles figurant déjà dans la preuve de Janet) , en le traduisant en un problème de prolongement de telles immersions.

Ce problème a été en fait considéré par Darboux [5] dans le cas de surfaces. Étant donnée une courbe H contenue dans une surface M, et analytiquement isométriquement plongée dans $\mathfrak{R}^3$, à quelle condition ce plongement isométrique se prolonge dans un voisinage de H dans M ? La formule de Gauss , montre que si un tel prolongement existe, alors nécessairement, H est moins courbé dans M que dans $\mathfrak{R}^3$ (au sens de la courbure géodésique, pensez par exemple aux courbes géodésiques de M). Darboux démontre que cette condition est également suffisante. Jacobowitz dégage une généralisation de cette condition lorsque H est une hypersurface dans une variété riemannienne. C'est une condition de non-dégénérescence, à comparer avec la condition de liberté ci-dessous. Si cette condition est satisfaite, on arrive alors à écrire le problème de prolongement, sous forme de système de Cauchy–Kowalewka standard.

11 Cas C^1 : Travaux de Nash (1954) et Kuiper (1955)

En 1954 [27], J. Nash, surprend le monde mathématique, en passant de l'hypothèse courante d'analyticité des immersions isométriques, à la régularité la plus "faible" imaginable, c'est-à-dire C^1. De plus, il passe de considérations locales au problème global. On peut dire que ce fût là la première révolution nashienne dans l'évolution de cette histoire d'immersions isométriques. Ses résultats ont été complétés par N . Kuiper en 1955 [22]. Les deux auteurs montrent ainsi le résultat "incroyable" :

Dès qu'une variété riemannienne (V^n, g) *admet une immersion* différentiable *dans* $\mathfrak{R}^q$*, alors, elle admet une immersion isométrique* C^1*, si* $q > n$.

(L'amélioration de Kuiper par rapport à Nash était de passer de $q > n + 1$ à $q > n$).

Deux exemples d'applications de ce résultat qui défient l'imagination sont : un tore plat de dimension 2, et le plan hyperbolique, isométriquement plongés dans $\mathfrak{R}^3$, ou même dans un ouvert quelconque de $\mathfrak{R}^3$! Ce résultat contredit aussi une vieille conjecture selon laquelle le théorème de rigidité de la sphère s'étend au cas C^1.

Pour "modérer" un peu "l'effet choquant" dans ce résultat, disons qu'à ce niveau de régularité, la courbure, qui est l'invariant infinitésimal principal des métriques riemanniennes, n'a pas de sens. Les obstructions usuelles aux immersions isométriques proviennent de la comparaison de la courbure intrinsèque avec la courbure ambiante, via la formule de Gauss .

12 h-principe de Gromov

La lettre h dans le terme h-principe désigne le mot homotopie. Le h-principe est valable pour un système d'équations aux dérivées partielles, dès que toute solution formelle peut être homotopée à une vraie solution (solution holonome). Le concept de solutions formelles a été déjà introduit par E. Cartan, et consiste à considérer les dérivées partielles comme des variables indépendantes.

Le h-principe permet d'avoir des solutions "bon marché" (suivant les termes de D. Sullivan). Il permet en particulier de montrer l'abondance de solutions, dès qu'il y en a une. M. Gromov, démontre le h-principe sous les hypothèses du théorème de Nash-Kuiper. Cela revient à dire, essentiellement que toute immersion C^1 est C^1-homotope à une immersion isométrique C^1. Plus spectaculaire (mais dans le même esprit), sous certaines conditions topologiques (assez faibles) ajoutées aux hypothèses de Nash-Kuiper,

Toute application continue $V^n \to W^q$, $q > n$ peut être approchée au sens de la topologie C^0 (fine dans le cas non-compact) par des immersions isométriques C^1 ([14], p 204) .

13 Cas C^k : Grand Théorème de Nash

En 1956 [28], J. Nash révolutionne une deuxième fois la théorie en montrant la possibilité de plongement lisse et global, d'une variété riemannienne quelconque dans un espace euclidien de dimension assez grande. Plus précisément, *soit (V, g) une variété riemannienne C^k, avec $3 \leq k \leq \infty$, alors, elle se plonge C^k isométriquement dans $\mathfrak{R}^q$, dès que :*

$q =\geq \frac{1}{2}n(3n + 11)$, si V est compacte,

$q =\geq \frac{1}{2}n(3n + 11)(n + 1)$, si V est non-compacte.

Ce n'est que 10 ans plus tard [29], qu'il démontre le même résultat de plongement (global) dans le cas de variétés *analytiques compactes.*

13.1 Amélioration de $q(n)$

Gromov et Rohlin [13] et [12, 14] améliorent la preuve de Nash, et ainsi démontrent un résultat valable, dans le cas compact ou non, et pour $3 \leq k \leq \omega$, avec $q = n^2 + 10n + 3$. (C^ω signifie analytique). [3]

13.2 Cas C^2

Ainsi, on voit, essentiellement dû à Nash, la possibilité d'immersions C^1 ainsi que C^k, pour $3 \leq k \leq \omega$ dans un espace euclidien. On ne sait encore rien sur le cas C^2. En d'autres termes, on ne sait pas si une métrique C^2 peut être immergée de façon C^2 dans un certain espace euclidien !

13.3 Le h-principe de Gromov

Dans l'approche de Nash (et reprise par les autres auteurs), on commence par considérer des applications $f : V^n \to \mathfrak{R}^q$ qui sont libres. Cela veut dire que le plan osculateur, i.e. celui engendré par les dérivées partielles de premier et second ordre de f est de dimension maximale, i.e. $\frac{1}{2}n(n + 1) + n$. Cela assure en particulier que le système est sous-déterminé, et aussi l'absence de toute obstruction infinitésimale, de type courbure (par exemple dû a la formule de Gauss ou Godazzi, ou ce qui peut s'en déduire).

[3] Voir le site : www.math.princeton.edu/jfnj/, où Nash commente une lettre qu'il a reçu en 1998, stipulant que sa preuve donne des immersions qui ne sont pas nécessairement des plongements. De toute façon les améliorations apportées par d'autres auteurs remédient en particulier à ce problème.

Gromov démontre [14] le h-principe pour les applications libres isométriques, de V^n dans $\mathfrak{R}^q$ dès que $q \geq \frac{1}{2}(n+2)(n+3)$, et ce dès que la classe de différentiabilité k est ≥ 5 (jusqu'au cas analytique), et pour toute variété (compacte ou ouverte).

14 Théorème de la fonction implicite dans les espaces de Fréchet

L'article de 1956 de Nash [28] était écrit de manière "brutale", [4] ce qui a fait dire a M. Gromov, qu'il pense qu'il est le seul à l'avoir vraiment lu [23]

Mais, c'est en fait J. T. Schwartz qui a remarqué que cet article contient (implicitement) un théorème de fonction implicite sur les espaces de Fréchet [34]. Ceci n'a pas empêché d'autres mathématiciens d'essayer d'expliciter davantage, et d'utiliser dans d'autres domaines ce théorème. Citons à titre d'exemple : Moser, Kolmogorov, Arnold, Sternberg, Jacobowitz, Gromov, R. Hamilton ... Le théorème de la fonction implicite sur les espaces de Fréchet porte actuellement le nom du théorème de Nash-Moser (ou même Newton-Nash-Moser, car sa preuve si délicate repose sur la méthode d'itération de Newton). Ce théorème a été utilisé en particulier en théorie des systèmes dynamiques dans la preuve du célèbre théorème KAM (Kolmogorov-Arnold-Moser).

15 Cas pseudo-riemannien

Une variété (V, g) pseudo-riemannienne (ou semi-riemannienne, suivant d'autres terminologies, mais à ne pas confondre avec sous-riemannienne) est telle que g soit un champ de formes quadratiques non-dégénérées de signature (n_-, n_+), i.e. sur tout espace tangent la forme est isomorphe à la forme $-x_1^2 \ldots - x_{n_-}^2 + y_1^2 + \ldots + y_{n_+}^2$ sur $\mathfrak{R}^n$, et $n = n_- + n_+$. Un sous-espace tangent $E \subset T_xV$ est isotrope si la forme quadratique g_x y est identiquement nulle. La dimension maximale d'un tel E vaut exactement l'indice de g, $ind(g) = \min(n_-, n_+)$. Il est clair que si une immersion isométrique $(V, g) \to (W, h)$ existe, alors $ind(g) \leq ind(h)$. Moyennant cette condition, Gromov [12] et Greene [11], généralisent toute la théorie de Nash, avec des constantes adaptées (voir l'exposition de [14] qui contient des résultats supplémentaires).

En particulier toute variété pseudo-riemannienne de type (n_-, n_+) admet un plongement isométrique dans un espace de Minkowski $\mathfrak{R}^{q_-,q_+}$, i.e. l'espace plat $\mathfrak{R}^q$, $q = q_- + q_+$ muni partout de la forme :

$$-x_1^2 \ldots - x_{q_-}^2 + y_1^2 + \ldots + y_{q_+}^2$$

dès que $\min(n_-, n_+) \leq \min(q_-, q_+)$ *et q est assez grand.*

[4] La vie de Nash, y compris durant cette période ultra-créative, était très turbulente. On estime que son travail lui valait la médaille Fields, très probablement déjà en 1958, et certainement en 1962, mais les complications de son état de santé dans cette dernière période l'ont mis "hors jeu". Ce n'est qu'en 1994, qu'il obtient le prix Nobel d'économie, pour un travail sur la théorie de jeux, qu'il a effectué au début de sa carrière. Voir par exemple, le site www-gap.dcs.st-and.ac.uk/ history/Mathematicians/Nash.html pour une biographie assez dramatique sur la vie de Nash, ainsi que [24] et [26].

Dans le cas d'immersions isométriques C^1, Nash et Kuiper ont utilisé des applications courtes (i.e. lipschitziennes, à constante de Lipschitz 1) qu'ils rendent ensuite isométriques. Dans [14], on arrive à adapter cette notion non-évidente au cas pseudo-riemannien, et ainsi généraliser le théorème de Nash-Kuiper dans le cas C^1, à cette situation.

Mais Greene et Gromov n'étaient pas les seuls à étudier le problème pseudo-riemannien. On peut remarquer à ce propos :

- La preuve du théorème d'immersion local analytique, s'adapte directement au cas pseudo-riemannien, notamment suivant l'approche de Jacobovitz (en formulant le problème sous forme de prolongement).
- On savait que le problème pseudo-riemannien général contient le cas riemannien au sens suivant.Si l'on veut immerger isométriquement (V, g) dans (W, h), on construit $(V \times W, g \bigoplus -h)$, et on immerge dedans V de façon isotrope (pour être cohérent ici, il faudrait étendre la discussion, comme c'est fait dans [14], au cas des métriques pseudo-riemanniennes "dégénérées", i.e. des champs de tenseurs covariants symétriques d'ordre 2 quelconques).
- Le problème pseudo-riemannien est parfois plus facile que le cas riemannien. Ainsi le plan hyperbolique $\mathfrak{H}^2$ admet une réalisation connue sous le nom de modèle de l'hyperboloide dans l'espace de Minkowski $\mathfrak{R}^{1,2}$. On peut donc dire, d'une part que le théorème de Hilbert est faux si l'on remplace l'espace euclidien par l'espace de Minkowski. D'autre part, le théorème de rigidité de la sphère dans l'espace euclidien ne se généralise pas à cette situation. En effet, l'hyperboloide peut être vu comme une sphère de rayon imaginaire dans l'espace de Minkowski, et il se trouve qu'il admet des déformations isométriques non-triviales [17].

16 Obstructions en codimension relativement petite

16.1 Géodésibilité d'une sous-variété

On apprend au premier cours de géométrie "intrinsèque", à différencier entre les surfaces *planes* et les surfaces *plates* dans l'espace euclidien. Les premières sont contenues dans des plans et les secondes sont simplement à courbure de Gauss nulle. Des exemples de ces dernières sont les cônes et cylindres généralisés. Modulo une rotation de l'espace, ces derniers consistent à se donner une section s, c.à-d. une courbe de Jordan lisse horizontale, et ensuite de la saturer par les droites verticales. On voit sur cet exemple que même si le cylindre n'est pas plan, il l'est à un deuxième ordre. On veut dire par cela qu'il est engendré par des droites. C'est un fait général :

Une surface plate de l'espace euclidien est engendrée par des droites le long desquelles, son plan tangent est parallèle (voir [33] pour une approche moderne).

Notons cependant, qu'en dehors du cas analytique, une surface plate, peut contenir une partie (strictement) plane et une autre *réglée* comme on vient de le décrire. Dans le cas du cylindre ci-dessus, la partie plane correspond aux segments de droites contenus dans sa section s (c'est le pur hasard qui permet dans cette situation de prolonger les droites génératrices à la partie plane).

En général, le défaut de planitude (i.e. géodésibilité) d'une sous-variété V d'une variété

riemannienne W (tout étant C^∞ ici, pour simplifier), se mesure à l'aide de sa seconde forme fondamentale, une forme (vectorielle) bilinéaire symétrique π.

L'espace plat (dit aussi espace de nullité relative) Fl_x de V en x est le noyau de π_x :

$$Fl_x = \{X \in T_xV / \pi_x(X, Y) = 0, \forall\ Y \in T_xV\}$$

On a la propriété suivante d'intégrabilité dans le cas des sous-variétés des variétés à courbure *constante*, généralisant le constat ci-dessus sur les surfaces plates de l'espace euclidien (voir par exemple [14] pour une approche synthétique) :

Supposons la variété ambiante W à courbure constante. Alors, dans l'ouvert $\mathcal{U}$ de V où sa dimension est minimale, l'espace plat Fl de V est intégrable et ses feuilles sont géodésiques dans W (et à fortiori dans V). On l'appellera le feuilletage plat de V et notera $\mathcal{F}l$.

16.2 Lemme algébrique

[35] et ([14], p 262). *Soit $\pi : \mathfrak{R}^n \times \mathfrak{R}^n \to \mathfrak{R}^{q-n}$, une forme biliéaire symétrique vérifiant :*

$$\langle \pi(X, X), \pi(Y, Y) \rangle - \langle \pi(X, Y), \pi(X, Y) \rangle = 0, \forall X, Y \in \mathfrak{R}^n,$$

alors, la dimension du noyau de π est $\geq 2n - q$.

On continue à améliorer et généraliser ce résultat à des situations similaires, par exemple en remplaçant l'égalité $= 0$ par ≤ 0 (voir par exemple [12]).

16.3 Le feuilletage local $\mathcal{F}l$

La seconde forme fondamentale d'une immersion entre variétés à *même* courbure *constante* satisfait à la condition ci-dessus (grâce à l'équation de Gauss). On en déduit :

Soit $V^n \to W^q$ une immersion isométrique entre variétés de même courbure constante, alors, l'espace plat de V vérifie : $dim(\mathcal{F}l) \geq 2dim(V) - dim(W) = dim(V) - codim(V) = 2n - q$.

La difficulté majeure réside dans le fait que $\mathcal{F}l$ n'est défini que dans $\mathcal{U}$ qui est généralement un ouvert propre de V. Ailleurs, la dimension de Fl peut exploser. Heureusement, on arrive a montrer qu'une telle explosion ne se produit pas en suivant une feuille de $\mathcal{F}l$ dans $\mathcal{U}$. En d'autres termes, les feuilles de $\mathcal{F}l$ sont *relativement complètes* dans V, ce qui signifie lorsque V est suppposée complète, que les feuilles de $\mathcal{F}l$ le sont également [7, 14].

Toutes ces informations algébrico-géométriques servent à démontrer le résultat :

Il n'existe pas d'immersion isométrique d'un tore plat de dimension n dans $\mathfrak{R}^{2n-1}$ [30],

et surtout le résultat de rigidité suivant :

Soit $f : V^n \to W^q$ une immersion isométrique, où V et W ont une même courbure constante non-nulle. *Si, V est compacte et $2n - q > 0$, alors $f(V)$ est totalement géodésique dans W (voir [8] (1975) et une autre jolie preuve dans ([14], p) en courbure positive, et [37] (1985) en courbure négative*).

16.4 Intervalles

On observe ainsi que pour les immersions isométriques $V^n \to W^q$:

• Il y a une riche (et encore ouverte) théorie d'obstructions, si $q < 2n$.

• Pour $q \geq n(n+1)/2$, des immersions isométriques locales existent en général (Cartan-Janet dans le cas analytique et [20] dans le cas lisse).

• Pour $q \geq q(n)$, où $q(n)$ est le polynome donné par Nash (ou celui amélioré par Gromov), les immersions isométriques (globales) existent et abondent.

• Pour $2n \leq q \leq q(n)$, on ne sait rien (sur les immersions globales) !

17 Coup de théâtre : Preuve de Günther

Autour de 1987, vivant en Allemagne de l'est, M. Gunther, trouve une démonstration du théorème des immersions isométriques, cas lisse, ($C^{2,\alpha}$, $\alpha > 0$), par une méthode classique, c'est-à-dire en d'autres termes, en utilisant plutôt un théorème de fonction implicite classique sur les espaces de Banach (ou disons plus précisément le principe d'applications contractantes standard). Sa preuve occupe quelques petites pages (voir [15, 16] et les expositions [1, 36]).

Décidément, Gunther n'était guère impressionné par la réputation de la difficulté singulière du problème, ou peut-être simplement les échos de cette réputation ne lui sont-ils pas parvenus (à cause du mur de Berlin ?). C'est dire que l'adage "un homme averti en vaut deux" est "un théorème" faux dans la recherche mathématique (et probablement dans la recherche scientifique en général). Mais ce n'est pas par cette "sagesse" que nous voudrions terminer notre histoire, car, d'une part, cette histoire reste encore inachevée, et d'autre part, qu'on ne se trompe pas, le théorème de Nash-Moser reste d'un intérêt exceptionnel en mathématiques, même si son usage reste très délicat. Ceci conduit à se poser la question, peut-on touver des preuves classiques, à des théorèmes célèbres (e.g. KAM) abordés actuellement par la méthode de Nash ? Je remercie F. Béguin et T. Barbot pour leurs remarques et suggestions.

Bibliographie

[1] S. Alinhac, P. Gérard, *Opérateurs pseudo-différentiels et théorème de Nash-Moser. Savoirs Actuels. InterEditions*, Éditions du CNRS, Paris (1991).

[2] E. Cartan, *Sur la possibilité de plonger un espace riemannien donné dans un espace euclidien.* Ann. Soc. Pol. Math. 6 (1927) 1-7.

[3] S. Chern, *A proof of the uniqueness of Minkowski's problem for convex surfaces*, Amer. J. Math. 79 (1957) 949–950.

[4] S. Cohn-Vossen, *unstarre gescholassen flächen*, Math. Ann. 102 (1929) 10-29.

[5] G. Darboux, *Leçons sur la théorie générale des surfaces*, Vol. 3. Gauthier-Villars, Paris (1984)

[6] V. Efimov, *Generation of singularites on surfaces of negative curvature*, (Russian) Mat. Sb. (N.S.) 64, no. 106 (1964) 286–320.

[7] D. Ferus, *On the completeness of nullity foliations*, Michigan Math. J. 18 (1971) 61–64.

[8] D. Ferus, *Isometric immersions of constant curvature manifolds*, Math. Ann. 217, no. 2 (1975) 155–156.

[9] L. Florit, *On submanifolds with nonpositive extrinsic curvature.*, Math. Ann. 298, 1, (1994) 187–192.

[10] J. Gasqui, *Sur l'existence d'immersions isomtriques locales pour les varits riemanniennes*, J. Differential Geometry 10 (1975) 61–84.

[11] R. Greene, *Isometric embeddings of Riemannian and pseudo-Riemannian manifolds*, Memoirs of the American Mathematical Society, No. 97 (1970).

[12] M. Gromov, *Isometric imbeddings and immersions*, Dokl. Akad. Nauk SSSR 192 (1970) 1206–1209.

[13] M. Gromov, V. Rohlin, *Imbeddings and immersions in Riemannian geometry*, Uspehi Mat. Nauk 25 no. 5 (155) (1970) 3–62.

[14] M. Gromov, *Partial differential relations* Springer-Verlag, Berlin (1986).

[15] M. Günther, *On the perturbation problem associated to isometric embeddings of Riemannian manifolds*, Ann. Global Anal. Geom. 7, n. 1 (1989) 69–77.

[16] M. Günther, *Isometric embeddings of Riemannian manifolds*, Proceedings of the International Congress of Mathematicians, Vol. I, II (Kyoto, 1990), Math. Soc. Japan, Tokyo (1991) 1137–1143.

[17] J. Hano, Jun-ichi, K. Nomizu, *On isometric immersions of the hyperbolic plane into the Lorentz-Minkowski space and the Monge-Ampre equation of a certain type*, Math. Ann. 262, no. 2(1983) 245–253.

[18] D. Hilbert, *Uber Flächer von konstanter Gaußscher Krummung*, Tans. AMS 2 (1901) 87-99.

[19] H. Jacobowitz, *Local isometric embeddings. Seminar on Differential Geometry*, pp. 381–393, Ann. of Math. Stud., 102, Princeton Univ. Press, (1982) Princeton, N.J.

[20] H. Jacobowitz, *Extending isometric embeddings*, J. Differential Geometry 9 (1974), 291–307.

[21] M. Janet, *Sur la possibilité de plonger un espace riemannien donné dans une espace euclidien*, Ann. Soc. Pol. Math. 5 (1927) 38-43.

[22] N. Kuiper, *On C^1-isometric imbeddings. I, II*, Nederl. Akad. Wetensch. Proc. Ser. A. 58 = Indag. Math. 17 (1955), 545–556, 683–689.

[23] R. Langevin, *Interview with Mikhael Gromov*, Development of mathematics 1950–2000 (2000) 1213–1227, Birkhäuser, Basel.

[24] J. Milnor, *John Nash, and A beautiful mind*, Notices Amer. Math. Soc. 45, no. 10, (1998) 1329–1332.

[25] T. Milnor, *Efimov's theorem about complete immersed surfaces of negative curvature*, Advances in Math. 8 (1972) 474–543.

[26] S. Nasar, *A beautiful mind*, Simon and Schuster (1998) New York.

[27] J. Nash, *C^1 isometric imbeddings*, Ann. of Math. (2) 60 (1954) 383–396.

[28] J. Nash, *The imbedding problem for Riemannian manifolds*, Ann. of Math. (2) 63 (1956) 20–63.

[29] J. Nash, *Analyticity of the solutions of implicit function problems with analytic data*, Ann. of Math. (2) 84 (1966) 345–355.

[30] B. O'Neill, *Isometric immersion of flat Riemannian manifolds in Euclidean space*, Michigan Math. J. 9 (1962) 199–205.

[31] R. Sacksteder, *The rigidity of hypersurfaces*, J. Math. Mech. 11 (1962) 929–939.

[32] J-M. Schlenker, *Complete surfaces with negative extrinsic curvature*, Preprint.

[33] M. Spivak : *"A comprehensive introduction to differential geometry"*, Berkely, Publish or Perish (1979).

[34] J. Schwartz, *On Nash's implicit functional theorem*, Comm. Pure Appl. Math. 13 (1960) 509–530.

[35] C. Tompkins,*Isometric imbedding of flat manifolds in Euclidean space*, Duke Math. J. 5 (1939) 38-61.

[36] D. Yang, *Günther's proof of Nash's isometric embedding theorem*, xxx.arXiv.cornell.edu/pdf/math.DG/9807169

[37] A. Zeghib, *Feuilletages géodésiques appliqués*, Math. Ann. 298, no. 4 (1994) 729–759.

150 ans de Géométrie Riemannienne

Marcel Berger

1 Introduction

Ce texte ne comporte pratiquement aucune affirmation de nature épistémologique, ou de nature historique. Nous espérons cependant qu'il apportera indirectement une contribution tant à l'histoire de la géométrie qu'à son épistémologie. En effet depuis la création de la géométrie riemannienne, le volume des résultats obtenus dans cette discipline, est énorme, et il est donc difficile de s'y orienter, même pour un spécialiste de cette discipline. Dans ce panorama, nous donnerons l'état de la question (au mieux de nos connaissances) aujourd'hui, en omettant complètement, sauf dans quelques cas très particuliers, les résultats intermédiaires. Nous espérons cependant que la vision simplifiée de cette géométrie, que nous proposons, pourra rendre des services.

Définissons notre sujet ; il ne s'agit pas de la géométrie différentielle globale, ni même des structures variées que l'on peut attacher à une variété riemannienne. Par exemple, il ne sera pas question de spineurs, de champs de Yang-Mills, ni du théorème de l'indice ou de K-théorie. Enfin, la plupart du temps, nous avons fait l'impasse sur la géométrie riemannienne des dimensions 3 et 4. Dans ces deux dimensions en effet, la géométrie riemannienne est, à de nombreux égards, beaucoup plus riche, des résultats récents y ont été obtenus, d'autres spectaculaires s'annoncent. Elle contribue en effet de façon importante à la connaissance de la topologie différentielle dans ces dimensions.

Nous resterons vraiment terre à terre, sur la variété elle-même, à la seule exception de la section VI où nous grimperons sur son fibré tangent. Une grande partie de ce que nous dirons correspond à ce qui est appelé "coarse geometry" par Hitchin. [23][1].

Nous adopterons la terminologie contemporaine, omettant les étapes intermédiaires. Nous ne donnerons pratiquement aucune référence, aucun nom, renvoyant le lecteur pour références et détails systématiquement à nos article [2], [5], sauf pour des références plus récentes que nous donnons alors explicitement. Une raison historique, pour laquelle il est

[1] un très bon texte de résumé et de prospective sur la géométrie différentielle. Le panorama proposé restera pourtant, pour des raisons de place, partiel, voire très partiel et personnel. Nous prions tous ceux que nous n'avons pas cités de bien vouloir nous en excuser.

difficile de citer de noms, est le fait la plupart des résultats présentés sont l'œuvre de plusieurs, et non celle d'une personne unique. Pour plus de précisions, on pourra consulter aussi la collection des tomes de géométrie de l'encyclopédie de chez Springer, mais les informations y sont dispersées.

Nous mentionnerons bien sûr de nombreuses questions qui restent ouvertes. Nous ne les hiérarchiserons pas, d'une part par manque de don prophétique, mais aussi au nom de la note épistémologique empruntée à [17] (p. 130) : "quand on dit qu'une question est naturelle, le mot naturel peut mener dans deux directions complètement différentes : celle résumée par $i-e-s$ (intellectual-emotional-social), et celle résumée par "$m-l-p$" (mathematics-logic-philosophy), et la première ne conduit pas toujours à la deuxième." Seule l'histoire ou les génies peuvent faire de telles connexions. Les problèmes ouverts que nous mentionnerons ne sont donc pas tous importants [2].

2 L'héritage de Riemann en langage moderne

En 1854, l'effort de Riemann pour fonder la géométrie, nous léguait la notion de variété riemannienne (VR). Il s'agissait d'une généralisation de la géométrie euclidienne. En outre, il devenait possible de donner une base solide à la géométrie hyperbolique alors en phase de consolidation (et ce de surcroît en toute dimension). Dans ce travail, l'intetion de Riemann était d'étendre la géométrie, et plus particulièrement la géométrie des surfaces de Gauss , mais il pensait aussi certainement au problème de la représentation conforme pour les "surfaces de Riemann" compactes, nécessairement abstraites.

Une VR est un espace (lisse) dont les voisinages de chaque point, lorsqu'ils sont "zoomés" infiniment, deviennent des espaces euclidiens, qui sont les espaces tangents en ces points. On notera $(M^n; g)$ une telle variété riemannienne, T_pM son espace tangent en un point p et $g(p)$ désignera donc la métrique euclidienne sur T_pM.

Une VR hérite automatiquement de quantité de données à contenu géométrique : des angles pour les vecteurs tangents (en un même point), une longueur pour toute courbe c, une métrique obtenue $d(p,q)$ comme borne inférieure de la longueur des courbes joignant p à q (qui fournit d'ailleurs la même topologie que celle initiale de variété), et c'est par rapport à cette topologie que nous parlerons de "complétude" et de "compacité". Il est de plus associé à la structure riemannienne, une mesure dp, qui est en quelque sorte l'intégrale des mesures de Lebesgue des T_pM associées aux métriques $g(p)$. Toute sous-variété hérite d'une structure riemannienne, donc aussi d'une mesure canonique. Joue enfin un rôle déterminant la notion de géodésique : ce sont les courbes qui réalisent entre deux de leurs points assez voisins la distance de ces points (on parle alors de segment géodésique ou de plus courbe géodésique). Mais si on considère une géodésique globalement, elle va cesser en général (et toujours dans le cas compact) d'être un segment géodésique. Une des raisons pour introduire plus bas des invariants comme la courbure est que, si les géodésiques sont solutions d'équations déductibles de g, en revanche les distances ne le sont pas en général.

[2] D'autres références à contenu plus épistémologique sont : [1], [4], [16], [17], [23], [25], [34]. La référence [17], à laquelle nous empruntons beaucoup dans la dernière section, est extrêmement dense, comme le sont souvent les textes de cet auteur, mais constitue un ensemble de plusieurs programmes de recherche pour le siècle qui vient de s'ouvrir.

Dans toute la suite, une VR sera supposée compacte ou au moins complète. Le cas intermédiaire extrêmement intéressant (par exemple pour les applications à la théorie des nombres) est celui des variétés de volume fini, mais nous ne pourrons pas en parler ici par manque de place.

Questions :

Des questions géométriques naturelles (!) sont de calculer, ou d'estimer en tout cas, le troisième côté d'un triangle : $a = d(q,r) = F(b,c,\alpha)$ en fonction des longueurs de deux autres côtés (b,c) et de l'angle qu'ils forment, ou encore, ce qui revient pratiquement au même (mais à un titre plus faible), à étudier la somme des angles d'un triangle. On peut, on doit (voir plus bas) aussi, se poser la question de ce que devient dans les VR l'inégalité isopérimétrique euclidienne classique.

Que savait (exprimé dans les termes de la topologie moderne) Riemann ? Il introduisait la courbure sectionnelle K qu'il définissait ainsi : si on regarde de près le développement limité de la longueur d'un arc de cercle de rayon α construit avec deux vecteurs unitaires tangents (x,y) formant l'angle α on remarque que le premier terme est celui euclidien, ensuite il n'y a jamais de terme d'ordre 2 (une VR est toujours symétrique, à l'ordre deux (voir plus bas), le terme d'ordre 3 s'écrit $K(x,y)$, soit plus précisément :

$$\text{longueur} = \alpha\varepsilon\left(1 - \tfrac{\mathrm{K}(x,y)}{6}\varepsilon^2 + o\left(\varepsilon^2\right)\right)$$

Le réel $K(x,y)$ ne dépend en fait que du plan tangent P engendré par x et y, et fournit donc une application K : Grass

$$(M) \to \mathbb{R}$$

de la grassmannienne des plans (de dimension 2) tangents à M^n. La donnée de la courbure sectionnelle est équivalente en fait à celle du tenseur de courbure R de Riemann classique de type $(3;1)$. Elle dépend donc comme lui de $\frac{n^2\left(n^2-1\right)}{12}$ paramètres. Pour les surfaces, elle s'identifie à la courbure de Gauss ; en dimension 3 elle dépend de 6 paramètres, soit autant que la métrique, mais à partir de $n = 4, 5, ...$ elle dépend d'un nombre de paramètres proche de celui des dérivées secondes de la métrique, soit $\frac{n^2(n+1)^2}{4}$. On peut donc espérer reconstruire la métrique g, de même que l'on reconstruit une fonction par une double intégration à partir de ses dérivées secondes (par exemple la positivité entraîne la convexité, ou permet des contrôles plus généraux comme en mécanique hamiltonnienne, etc.)

Riemann savait que la courbure K mesure très bien le défaut des VR d'être des espaces euclidiens : si K est identiquement nulle, alors notre VR est localement isométrique à l'espace euclien E^n. Mais on sait en fait plus : si la courbure K est une constante k, alors la variété est partout isométrique localement avec la sphère de rayon $\frac{1}{\sqrt{k}}$ si $k > 0$ et l'espace hyperbolique $Hyp^n(\mathrm{k})$ de courbure constante k, si k est négatif. On réunira ces trois espaces modèles en une seule notation $S^n(k)$, couvrant tous les réels k. Si de plus la VR considérée est simplement connexe (i.e. $\pi_1 = 0$), alors il y a isométrie globale (et donc unicité de la métrique à difféomorphisme près)[3].

[3] Juste avant 1920, on disposait avec les travaux de Ricci et Levi-Civita d'un calcul différentiel intrinsèque de tout ordre, et de la notion de transport parallèle. Einstein fut bienheureux de trouver cela pour développer la Relativité générale. Mais cette exception et de la "conjecture de la masse positive" (voir [22], chapitre 6), il nous semble que les liens entre la Relativité générale et la géométrie riemannienne, entendue au sens strict, sont pratiquement inexistants (voir la figure 0.43 de [6]).

3 Formes d'espace, espaces localement symétriques (Riemann plus Klein)

3.1 Espaces 3-transitifs

En géométrie euclidienne, le cas d'égalité des triangles, c'est-à-dire que le groupe des isométries de E^n est transitif sur les triplets de points ayant les mêmes distances deux à deux exprime une propriété de 3-transitivité. Dans l'esprit de Klein, on peut demander quelles sont les VR qui ont cette propriété (locale, puis globale). Localement, cette 3-transitivité est équivalente à la constance de la courbure K ; de telles VR sont donc localement isométriques à l'une des $S^n(k)$. S'il y a simple connexité, il y a donc unicité. Mais le problème posé dès les années 1900, dit des formes d'espace de Clifford-Klein, consiste à décrire toutes les VR compactes à courbure constante, donc les quotients compacts par un groupe convenable d'isométries des $S^n(k)$. Pour les cas $k = 0$ et $k = 1$ (correspondant après normalisation à $k > 0$) la question est pratiquement terminée, la classification est connue, à quelques difficultés algorithmiques près [41] : dans le cas $k = 1$, il faut classer les groupes finis convenables des groupes orthogonaux ; dans le cas $k = 0$, il faut trouver les groupes finis convenables d'isométries des tores plats.

Le cas $k = -1$ pose des problèmes d'une toute autre ampleur ; sa résolution n'est pas du tout terminée, sauf pour $n = 2$ grâce à la représentation conforme. Dès $n \geq 3$ il est difficile déjà de trouver des exemples ; en toute dimension les premiers se trouvent dans l'article de [7]. Mais ils sont, pour le dire rapidement, construits de façon arithmétique, c'est-à-dire "en gros" donnés par des matrices convenables à coefficients entiers. On sait maintenant construire des exemples non arithmétiques [20], mais on n'a pas encore de classification systématique.

3.2 Espaces 2-transitifs

On peut se poser la même question pour les espaces de Riemann-Klein 2-transitifs, c'est-à-dire ceux dont les groupes d'isométries agissent transitivement sur les paires de points à égales distance. La structure locale est connue seulement depuis les travaux de Jacques Tits [39] pour le cas général, après ceux de Wang en 1952 pour le cas de courbure positive ; ces géométries sont exactement celles des espaces symétriques de rang un (voir plus bas), à savoir (outre les $S^n(k)$), celles des espaces projectifs CP^n, HP^n, CaP^2 sur le corps C des complexes, celui H des quaternions, et celle du plan projectif des octaves de Cayley, ainsi que leurs versions hyperboliques $Hyp^n_{C,H,Ca}$. Il faut savoir que leur courbure K décrit complètement l'intervalle $[1/4, 1]$ dans le cas positif et l'intervalle $[-1, -1/4]$ dans le cas négatif. Dans le cas simplement connexe, il n'y a que ces espaces.

L'existence de quotients compacts est triviale dans le cas positif, car le groupe fondamental ne peut être que l'identité ou $\mathbb{Z}_2$. Dans le cas de courbure négative, on trouve dans tous les cas des quotients compacts arithmétiques (voir encore une fois Borel, 1963)et la classification n'est terminée que dans les cas $\mathbb{H}$ et $\mathbb{C}a$, où l'on sait qu'il n'y a que des quotients arithmétiques, [21]. Pour les Hyp^2_C il existe des quotients non arithmétiques, [31], mais la question de la classification complète est ouverte ainsi que la question des plus grands n.

3.3 Espaces symétriques (localement)

La théorie des espaces symétriques peit être considérée comme le premier miracle de la géométrie riemannienne, en fait comme un nœud de forte densité dans l'arbre de toutes les mathématiques. Dans l'esprit de Klein, on peut chercher les VR telles que (localement) la symétrie géodésique (suffisamment localisée) autour de tout point soit une isométrie. En effet, par composition autour de points différents on obtient alors de suite l'homogénéité locale de l'espace. Ceci entraîne que le tenseur de courbure R est invariant par transport parallèle, car, l'invaraince de ce tenseur par la symétrie entraime que sa dérivée covariante DR vérifiera $(-1)^5 DR = -DR$ et est donc identiquement nulle. On doit à Elie Cartan dans les années 1926 d'avoir découvert que ces géométries sont, dans une dimension donnée, en nombre fini, et en outre toutes classées. De plus, les composantes irréductibles et simplement connexes des espaces obtenus sont tous des espaces que l'on rencontre en géométrie et en analyse : les grassmanniennes sur les corps variés ci-dessus, leurs pendants à courbure négative (donc non compacts), l'espace des structures complexes, des formes quadratiques, sur les $\mathbb{R}^n$, etc. avec en plus quelques exceptions en basses dimensions liées aux groupes de Lie exceptionnels. Ces espaces, dans le cas non compact, sont des objets privilégiés pour l'analyse harmonique depuis les travaux d'Harish-Chandra, car ils mélangent le compact (spectre discret) et le non compact **(spectre continu)** (voir par exemple les travaux de [40]). Les espaces irréductibles sont à courbure $K > 0$ ou $K < 0$, les parties plates étant des tores maximaux dans le cas positif ou sous-espaces vectoriels dans le cas négatif. Le rang est la dimension commune à toutes ces parties plates ; le cas de rang un est celui des espaces doublement transitifs rencontrés plus haut.

La classification des formes compactes (de rang plus grand que un) est terminée : leur existence (arithmétique) est montrée dans l'article de[7], mais depuis [30], on sait qu'il n'y en a pas d'autres. Il convient de noter - c'est important pour l'ensemble des mathématiques- que dans l'étude des formes d'espaces généralisées, le quotient compact est certes intéressant géométriquement, mais l'intérêt essentiel réside dans les groupes discrets qui leur donnent naissance, et on est donc ainsi transporté dans la théorie des nombres.

3.4 Espaces homogènes

Les espaces homogènes de groupes de Lie G/H (où H est un sous-groupe compact), justiciables donc d'une géométrie "à la Klein", sont certes intéressants : ils servent à fabriquer des exemples de types variés, mais on ne voit pas émerger de classification raisonnable, ils sont trop généraux si on n'en demande pas plus. Ce n'est pas dire qu'il n'y ait pas de problèmes fondamentaux, par exemple voir la section 0.8. A de [13] : il est ahurissant que l'on ne comprenne encore pas du tout la géométrie homogène non triviale la plus simple, à savoir celle des métriques invariantes à gauche sur la sphère S^3, qui est le groupe simple SU(2). On en sait encore moins sur les métriques homogènes de l'espace projectif RP^3, qui ne sont autres que les quotients de celles ci-dessus sur S^3. Ces métriques sont pourtant celles qui donnent le mouvement d'un corps solide autour de son centre de gravité (dit mouvement "à la Poinsot").

4 Programme de H. Hopf I : courbure et topologie

4.1 Introduction

Étudier les VR générales $(M^n; g)$ est impossible : quelques exemples montrent en effet que tout invariant raisonnable est incontrôlable. Le premier est le rayon d'injectivité $\mathrm{Inj}(g)$, c'est-à-dire la borne supérieure des réels i tels que toute paire de points dont la distance est inférieure à i est jointe par un unique segment. On peut encore en donner une autre définition : le rayon d'injectivité est la borne inférieure des j tels que toute paire de points dont la distance est supérieure à j est jointe par au moins deux segments différents. Il est toujours positif (strictement) si la variété est compacte, ce qu'on supposera dans toute cette section.

Les quatre contre-exemples de base au contrôle du rayon d'injectivité sont : les ellipsoïdes de révolution très aplatis, une haltère très serrée, et les rectangles plats de côtés $(\varepsilon, 1/\varepsilon)$ ou $(\varepsilon, 1)$. Cela montre que pour le contrôler, il faut contrôler la courbure dans les deux sens ainsi que le volume et le diamètre. Le lecteur se convaincra vite qu'une inégalité **isopérimétrique** est aussi sans espoir. Nous verrons en section 5 un contre-exemple sur le flot géodésique.

Heinz Hopf dans [24], proposait un double programme :

I. contrôler la topologie avec des fonctions de la courbure (mais "quelle expression de la courbure choisir" demandait-il ?).

II. il demandait si l'on pouvait métriser de façon intéressante une variété compacte donnée.

Ses motivations étaient multiples. D'abord en dimension 2 la formule $\int\limits_M K\,(p)\,dp = 2\pi\chi\left(M^2\right)$, dite théorème de Gauss -Bonnet[4],

(où χ est la caractéristique d'Euler de la surface) résolvait pratiquement le problème : en effet, on sait de suite quelles variétés peuvent porter des métriques de courbure de signe imposé. Heinz Hopf demandait par exemple si une condition de signe pour la courbure K d'une variété de dimension paire, impliquait que sa carctéristique $\chi\left(M^{paire}\right)$ devait avoir le signe évident. Ajoutons, pour les motivations de Hopf, qu'il savait démontrer ce résultat de constance du signe dans deux cas très particuliers : celui des formes d'espaces (espaces à courbure constante) et celui des hypersurfaces M^n des espaces euclidiens E^{n+1}. Il pensait aussi, bien sûr, à une généralisation en toute dimension paire de la formule ci-dessus.

Pour II, le théorème de la représentation conforme, en plus des travaux de Teichmüller, répondent entièrement à la question pour le cas de la dimension 2.

Le premier point est un paradigme de la problématique générale "passer du local au global" ; pour la courbure il s'agit plus exactement de passer "de l'infinitésimal au global". Outre son caractère intrinsèquement "naturel", cette situation se retrouve dans beaucoup de sujets de la physique : attraction newtonnienne, formation des cristaux, etc.

4.2 Courbure presque constante

Si la courbure est un aussi bon contrôle qu'on le croit, et puisque l'on connaît complètement (ou presque) les variétés compactes à courbure constante, alors on se doit de vérifier qu'une

[4] voir la section 0.E de [2] ou celle 3.4.5 de [5] pour l'histoire de cette formule)

(VR, g) dont la courbure varie très peu est en fait une variété qui porte une (autre) métrique à courbure constante. Une double surprise nous attend. Certes, cela est vrai dans le cas positif : il existe, en effet, pour tout n un $\varepsilon(n) > 0$ tel que l'hypothèse $1 - \varepsilon(n) < K < 1$ implique que M^n porte bien une métrique à courbure constante $+1$ [5].

Dans le cas de la courbure presque nulle, il faut normaliser le diamètre parce qu'en changeant d'échelle λ, la courbure varie comme λ^{-1}. Cela fait, il reste qu'il n'existe pas de $\varepsilon(n)$ comme ci-dessus car il existe d'autres variétés que les formes d'espace plates ; mais on les connaît toutes (à partir d'un certain $\varepsilon(n)$), ce sont les *infranil* variétés, quotients finis de quotients compacts de groupes de Lie nilpotents (1982).

Les choses sont encore plus surprenantes dans le cas négatif : pour tout $\varepsilon > 0$ et tout n > 4, il existe des (M^n, g) vérifiant $-1 < K < 1 - \varepsilon(n)$ pour lesquelles il n'existe aucune métrique à courbure constante. Mais la question ici n'est pas terminée.

4.3 Finitude et compacité

La dimension étant donnée, on considère en quelque sorte l'application qui à l'ensemble des VR compactes associe la variété support en oubliant la métrique. Si la courbure varie continûment, l'espace d'arrivée lui est discret, c'est celui des variétés différentiables de dimension donnée. On peut donc espérer que cette application vérifie une propriété de finitude à savoir que si la courbure varie dans un intervalle compact de $\mathbb{R}$, alors on ne trouve qu'un nombre fini de variétés possibles portant de telles métriques. Cela est faux et il existe beaucoup de contre-exemples : le lecteur pourra les imaginer à l'aide des figures données plus haut pour le rayon d'injectivité, mais il en existe aussi de plus sophistiqués. Cependant, on a bien depuis 1990, un tel théorème de finitude si $K > -1$, le diamètre est borné supérieurement et le volume borné inférieurement. On a aussi un théorème de finitude pour les VR telles que $\frac{Vol(g)}{Inj^n(g)}$ soit borné (en dimension donnée). Cela montre bien l'importance du rayon d'injectivité que l'on va retrouver de suite.

On peut, en effet, demander plus, à savoir si dans une topologie convenable, on a pour les VR des résultats de convergence, de compacité. Tel est bien le cas (grâce à des résultats obtenus entre 1970 à 1990) sous les quatre conditions suivantes : $|K| < 1$, volume borné inférieurement, diamètre borné supérieurement. Mais attention : la convergence a lieu seulement vers une métrique de classe $C^{1+\alpha}$ pour tout $\alpha > 1$ (on ne peut pas faire mieux : considérer un cône en son sommet). Le fait essentiel dans ce résultat est que ces quatre bornes entraînent une borne inférieure positive (explicite) pour le rayon d'injectivité (cela remonte à 1970).

Tous les résultats ci-dessus sont très généraux, en quelque sorte conceptuels. Mais dans l'esprit de Hopf, on pourrait demander tout simplement s'il existe des formules fournissant des invariants topologiques, comme les nombres de Betti, à l'aide de formules intégrales ayant pour intégrand un polynôme en le tenseur de courbure par exemple. On ne sait en gros que deux choses.

D'abord depuis 1943 et 1946, il existe en effet des formules (dues à Shiing Shen Chern) pour calculer la caractéristique d'Euler-Poincaré χ et les nombres caractéristiques (formés

[5] (voir [2] pour plus de détails, les premiers résultats complets ayant été obtenus aux alentours de 1970). La même référence de synthèse [2] pourra être consultée pour les détails et les références des deux cas suivants, ainsi que pour la plupart de tout ce qui suit dans ce texte, sauf autre référence explicite.

avec les classes de Pontryagin). Malheureusement, dès $n > 4$, ces formules ne peuvent pas être a priori bien contrôlées avec des inégalités sur la courbure ; on a affaire à des polynômes compliqués en la courbure et il semble difficile d'y faire apparaître la courbure sectionnelle. Cela à trois exceptions près : si $n = 4$, on sait que la constance du signe de la courbure entraîne la positivité de χ (1955). Puis, dans le cas kählérien, la première classe de Chern est donnée par la courbure de Ricci. Enfin les formules de Chern pour les nombres caractéristiques sont utilisées techniquement (mais sans valeur précise) pour remplacer la condition de borne inférieure sur le volume dans les résultats de compacité par la non-nullité "d'un nombre caractéristique au moins". En effet, quelle que soit la complexité du polynôme P(R), si $\int\limits_{M^n} P(R)\, dp \neq 0$, alors $|K| < 1$ entraîne que "Vol(g) est positif".

Enfin depuis 1981, on sait pour les nombres de Betti sur n'importe quel corps F que

$$\sum_i b_i(M^n, F) < \bullet(n) \exp\left(\sqrt{-Inf(K)}\right). \text{ Diamètre,}$$

pour une constante $\bullet(n)$ universelle. En particulier, si $K \geq 0$, les nombres de Betti sont bornés. La constante $\bullet(n)$ optimale est inconnue (est-ce 2^n, si on pense aux tores ?)

4.4 Courbure de signe constant

Nous venons tout juste de rencontrer les VR à courbure positive. On peut chercher plus généralement quelles sont les variétés compactes qui peuvent porter des métriques à $K > 0$, resp. $K \geq 0$, $K < 0$, $K \leq 0$. On peut dire aussi que l'on cherche à classer les géométries (riemanniennes) sur- ou sous-euclidiennes (hypo- ou hyper-). On connait à leur propos des résultats remarquables. Par exemple, concernant la somme des angles (α, β, γ) d'un triangle (suffisamment peitit) et les longueurs (a, b, c) (les côtes de longueur b et c déterminant l'angle α) [voir II.1], il y a dans le cas local équivalence entre

$$\alpha + \beta + \gamma > \pi \quad et \quad a^2 < b^2 + c^2 - 2bc\, cos\alpha$$

si $K > 0$ et de même dans les autres cas en changeant le sens de l'inégalité et en considérant les inégalités strictes au sens larges. Ce résultat date de 1925 pour le cas négatif et de 1952 pour le cas positif. Globalement, ces inégalités restent vraies pour tout triangle dans le cas positif (1960), tandis que pour le cas négatif, il faut ajouter visiblement une hypothèse, à savoir que l'espace est simplement connexe (toujours 1925).

Malgré cette belle équivalence géométrique, on est loin de savoir classer les variétés cherchées. En effet, il est en effet extrêmement difficile, dans les deux cas, de fabriquer des exemples (compacts). Si $K > 0$ on ne connaît que les espaces symétriques de rang un rencontrés en II.2, plus quelques cas en dimensions 6, 7, 13.

Un outil essentiel dans grand nombre des résultats cités, est que les inégalités $K > k$ et $K < k$ fournissent pour (tous) les triangles des inégalités de comparaison analogues à celles du cas $k = 0$, et ce pour tout nombre k réel, avec les triangles analogues à ceux mais construits dans les espaces modèles $S^n(k)$.

Si $K < 0$, il faut d'abord rappeler que, depuis 1898 (1925), on sait le revêtement universel de la variété (M^n, g) est l'espace vectoriel R^n (muni de la métrique image réciproque, et on a donc affaire à un groupe discret d'isométries). En conséquence, "tout est dans le π_1 de la

variété M^n". Le point essentiel est que, dans le cas $\mathrm{K} < 0$, le groupe π_1 M est un groupe hyperbolique (en fait c'est l'étude des variétés à $K < 0$ qui a conduit Gromov à la définition de ces groupes). Mais on ne sait toujours pas s'il y a ou non une différence entre la classe de ces π_1 et celle des groupes hyperboliques, et si oui, quelle est-elle ? De très nombreux travaux passés et en cours étudient la différence entre $\mathrm{K} < 0$ et $\mathrm{K} \leq 0$; comment la lire sur le π_1 ? Voire sur le flot géodésique (section V), ce qui est plus géométrique[6].

4.5 Courbure de Ricci.

Ce n'est pas parce que tout n'est pas fini entre courbure (sectionnelle) et topologie que l'on abandonne l'idée de contrôler (plus ou moins) la topologie avec un invariant plus faible, puisque K dépend de tant de paramètres : $\frac{n^2\left(n^2-1\right)}{12}$ contre seulement $\frac{n(n+1)}{2}$ pour la métrique. Ricci a introduit une courbure qui porte son nom $Ricci(g)$: $Ricci(g)$ d'un vecteur tangent x est la moyenne des courbures sectionnelles des plans passant par x. Ricci(g) est une forme quadratique, donc possède $\frac{n(n+1)}{2}$ paramètres. Il faut savoir, contrairement à ce que Ricci espérait, que sa forme est sans analogie avec la seconde forme fondamentale des surfaces de E^3, des hypersurfaces des E^{n+1}. A ce jour les lignes intégrales des directions propres de Ricci(g) par rapport à g n'ont avec elles aucune interprétation géométrique. Mais nous allons voir que, surtout depuis 1996, certaines inégalités sur $Ricci(g)$ ont d'énormes conséquences, et peuvent être, dans certains cas, aussi fortes que des inégalités sur la courbure sectionnelle. Le lecteur analyste qui calculera $Ricci(g)$ avec les dérivées secondes de la métrique g verra que les premiers termes sont ceux d'un opérateur du type laplacien, ce qui permet d'augurer des résultats (pensez à la théorie des fonctions harmoniques).

De même que la courbure (sectionnelle) contrôle le développement limité des arcs de cercle, la courbure de Ricci contrôle le développement de l'angle solide ω, autour d'une direction x égal à $\omega\varepsilon\left(1-\frac{Ricci(x)}{3}\varepsilon^2+o\left(\varepsilon^2\right)\right)$.

Malheureusement, on ne connaît pas, dès $n > 4$, les variétés à courbure de Ricci constante (même identiquement nulle). Ceci laisse mal augurer pour atteindre notre but. Cependant, les variétés à Ricci$(g) > (n+1)\,k$, où la constante k est un réel quelconque vérifient deux propriétés. Cependant, même que l'on savait intégrer les inégalités sur K, on sait intégrer Ricci$(g) > (n+1)\,k$. Cela donne un contrôle complet du volume de toutes les boules B$(p\,;r)$, pour tout point p et tout rayon r (même grand), (1963). La seconde chose est que Ricci contrôle le laplacien de la fonction distance f le long des géodésiques γ (1971). On a en effet

$$(\Delta(f\circ\gamma))' + \frac{1}{n-1}\left(\Delta(f\circ\gamma)\right)^2 < -Ricci\left(\gamma'\right).$$

Ceci explique l'on sait depuis 1996 que l'inégalité Ricci$(g) > (n+1)\,k$ entraîne en gros pour les triangles (comparés à ceux des espaces modèles $S^n\,(k)$, voir la sous-section précédente) des inégalités du type L^1 ou L^2. On peut dire que pour les géodésiques de conditions initiales données, on a seulement des probabilités pour savoir où se trouvent leurs extrémités.

Grâce à ces propriétés, on peut déduire depuis 1996 qu'il existe un $\varepsilon\,(n) > 0$ tel que l'inégalité Inf(Ricci). Diamètre $> -\varepsilon\,(n)$ entraîne que le groupe fondamental est nilpotent.

[6] Pour la construction d'exemples, très chère toujours, voir les travaux de [19] et [18].

C'est un théorème surprenant par rapport à celui de III.1 : d'une part on supprime la nécessité d'avoir une borne supérieure sur la courbure (sectionnelle), et d'autre part, on remplace la borne inférieure sur la courbure sectionnelle par une sur la courbure de Ricci.

Faisons cependant doublement attention. Des exemples montrent que l'hypothèse Ricci$(g) > 0$ n'interdit pas des nombres de Betti aussi grands que l'on veut, mais surtout on sait depuis 1994 que Ricci < 0 est une condition qui ne contraint pas la topologie de M . Pire, on peut approcher toute VR, certes au sens C^0, par des variétés avec Ricci < 0 .

On verra enfin en section VI que Inf(Ricci) permet (1980) de contrôler l'inégalité isopérimétrique, ce qui permet d'encadrer complètement le spectre entre autres choses.

4.6 Courbure scalaire.

La courbure scalaire scal(g) d'une (M^n, g) est la trace, par rapport à g, de la courbure de Ricci. $Scal(p)$ est aussi la moyenne de toutes les courbures sectionnelles des plans tangents en p. Si la théorie des espaces symétriques est le premier miracle de la géométrie riemannienne, ce que l'on sait de la condition scal > 0 en est le second. En effet, au moins dans le cas simplement connexe, on connaît exactement les variétés compactes qui admettent une métrique à scal > 0. Ce miracle n'est pas complet car les trois questions suivantes restent complètement ouvertes.

D'abord on ne connaît aucune conséquence géométrique (du genre des figures ci-dessus) de la condition scal > 0. Ensuite on ne connaît aucune description géométrique des variétés en question, seulement des conditions algébriques restreignant leur topologie : on trouve certes dans cette classe des espaces homogènes, mais aussi tous ceux que l'on peut en déduire par certaines chirurgies ; mais tout cela n'en donne pas une vision géométrique. Enfin la démonstration de ce résultat peut être obtenue par deux techniques bien différentes. La première utilise la chirurgie dans un sens et aux spineurs dans l'autre, la seconde utilise la théorie géométrique de la mesure pour obtenir des hypersurfaces minima, puis une déformation conforme. A ce jour aucun lien conceptuel n'est connu entre ces deux méthodes.

5 Programme de Hopf II : métrisation, meilleure métrique :

Il s'agit de voir ce qui peut substituer en dimensions plus grandes des techniques spéciales à la situation en dimension deux. En effet, dans cette dimension, la représentation conforme fournit les structures privilégiées cherchées et de plus l'ensemble de ces structures (espace des modules) forme un espace de dimension finie (Teichmüller). En outre on sait depuis 1988 que l'on peut obtenir ces métriques à courbure constante en suivant dans l'espace des métriques le flot naturel fourni par la courbure (de Gauss ici).

Il est important de noter que, dans tout ce qui suit, nous excluons pratiquement toujours le cas des dimensions 3 et 4, car la situation est radicalement différente dans le cas de dimension supérieure. D'une part parce que cela nécessiterait un texte entier. Et puis de remarquables travaux sont en cours. Enfin dans ces dimensions, l'interaction de la topologie avec la géométrie riemannienne est extrêmement forte : on peut espérer que la géométrie riemannienne apportera bientôt des contributions essentielles à la topologie des variétés dans ces deux dimensions 3 et 4. Mentionnons seulement, les travaux de [11] pour la dimension 4

et pour la dimension 3, ceux d'Anderson en cours et de [26] où les variétés considérées sont compactes.

5.1 Métriques critiques pour une fonctionnelle

Dès 1915, dans le cadre de la théorie de la relativité générale, donc pour des métriques lorentziennes (mais c'est valable aussi dans le cas riemannien), Hilbert introduisait la fonctionnelle utilisant la courbure la plus simple possible à savoir

$$F(g) = \int_M scal(g)\, dp,$$

normalisée avec par exemple $Vol(g) = 1$. Les métriques g critiques sont alors celles qui vérifent Ricci$(g) = kg$ (pour une constante k) dites "métriques d'Einstein." En dimension deux, nous ne sommes pas intéressés car la courbure de Ricci s'identifie à la courbure de Gauss , et la condition est vide ; en dimension trois la condition implique la constance de la courbure sectionnelle. En dimension quatre, la condition est encore assez forte, et par exemple on a des conditions topologiques pour l'existence sur une M^4, d'une structure d'Einstein[7]. En dimension $n \geq 5$ on ne connaît aucune restriction topologique nécessaire pour l'existence d'une telle métrique, et des travaux comme ceux de [32] semblent montrer que la situation est probablement indescriptible, du type de celles que nous allons voir plus bas. L'idée naturelle, qui marche comme on l'a vu en dimension 2, serait de suivre le flot de Ricci dans l'espace des métriques, c'est-à-dire d'étudier l'équation d'évolution

$$\frac{\partial g(t)}{\partial t} = -(\bullet)\, Ricci(g(t)),$$

avec une normalisation convenable. En dimensions 2, 3, 4 ce flot converge, dans certains cas, vers une métrique[8]. Mais en dimension 5, même s'il y a convergence pour t infini, on ne sait pas vers quoi.

On sait seulement (1987) que sur une variété donnée, l'ensemble des k possibles est discret et pour chaque k, l'espace des modules est de dimension finie et constitue un espace analytique réel stratifié[9].

Il ne faut pas croire que l'on en sait beaucoup plus dans le cas où la courbure de Ricci est nulle ; nous allons voir que l'on dispose d'exemples, mais ils sont tous sur des VR admettant des structures supplémentaires invariantes par transport parallèle. En existe-t-il d'autres ?

La fonctionnelle de Hilbert étant désespérante, puisque le flot de Ricci ne semble pas converger en général, on peut cependant utiliser le résultat de compacité cité en 4.3, c'est-à-dire étudier la fonctionnelle Diamètre(g) sur l'ensemble des métriques riemanniennes d'une variété M^n qui vérifient $|K| < 1$ et Vol(g)=1. Or le panorama est affreux, de nature fractale, ceci quelle que soit M^n, même pour la sphère. [33] montre que, dès que D est assez grand, l'ensemble $\{g : Diam(g) < D\}$ comprend un nombre exponentiel de composantes connexes, sur lesquelles on trouve une métrique optimale (minimum local, mais de classe $C^{1+\alpha}$). En

[7] Pour plus de détails, voir les références mentionnées ci-dessus et [2].
[8] voir plusieurs cas dans [2].
[9] Pour des résultats partiels dans certains cas, voir [2].

outre, ces minima sont exponentiellement voisins les uns des autres, un calcul sur ordinateur qui discrétiserait les données ne pourrait pas s'y retrouver pour choisir le minimum absolu. Pire, pour passer dans l'espace de toutes les métriques d'une composante connexe à une autre, la longueur d'un chemin est toujours une fonction non calculable de leur distance. Reste donc à trouver, si elles existent, des fonctionnelles inédites convenables.

Il existe cependant une catégorie (certes très particulière) de variétés pour lesquelles les choses marchent bien : ce sont les formes d'espaces à courbure négative. On sait depuis 1991 que sur elles, les métriques à courbure constante (voire localement symétriques dans le cas de courbure négative ou nulle) sont des minima absolu pour plusieurs fonctionnelles : le minimum de l'entropie du flot géodésique, le minimum de la fonction de comptage des géodésiques périodiques, et d'autres (pour ces deux notions "géodésiques" voir la section V). Et voici maintenant le troisième miracle (pas seulement) riemannien.

5.2 Le miracle kählérien

Entre les variétés localement symétriques, fort rigides certes, et les VR générales (même avec des conditions variées sur la courbure, etc.), il existe une catégorie intermédiaire, celle des variétés kählériennes (VK). Les VK sont les VR (M^n, g, J) où J est une structure complexe compatible avec g : invariante par transport parallèle (avec g elle-même bien sûr). Le miracle vient de ce que les différentes métriques associées à une J donnée sont exactement de la forme $g + \sqrt{-1}\partial'\partial'' f$, où f est une fonction sur M^n. On a donc affaire pour décrire le panorama des métriques kählériennes sur notre variété, non pas à un espace de formes différentielles quadratiques, très lourd, mais à un espace vectoriel de fonctions numériques. Reste à savoir ce que sont les différentes structures complexes J, mais elles forment toujours un espace de modules de dimension finie à difféomorphisme près.

Le panorama est parfait dans le cas où la première forme de Chern $c_1(M^n)$ est définie négative, nulle ou positive. On sait alors (1976-1977) que l'on peut suivre un flot sur l'espace des fonctions f qui converge vers une métrique d'Einstein, et en particulier si $c_1 = 0$, vers une métrique à courbure de Ricci nulle, d'où une famille de VR (VK) à courbure de Ricci nulle. Dans ce cas, le transport parallèle laisse une forme volume complexe, en outre invariante. Ce sont les variétés dites de Calabi-Yau, qui jouent aujourd'hui un rôle en physique théorique dans le cadre de la symétrie miroir.

Il existe une classe supplémentaire de VK encore plus spéciales, celles dites hyperkählériennes, où le transport parallèle laisse invariantes plusieurs structures complexes (un ensemble à deux paramètres).

La richesse kählérienne fait dire à certains que "la géométrie kählérienne est plus importante que la géométrie riemannienne".

5.3 Pas d'espoir d'autre miracle

Ne cherchez pas d'autres miracles du genre des espaces (localement) symétriques et des variétés kählériennes. En effet c'est un fait (1953) que les seules VR irréductibles qui admettent un invariant par transport parallèle autre que g elle-même (et sa forme volume) sont les espaces localement symétriques, les VK, les VK de Calabi-Yau, et les variétés hyperkählériennes.

Pour être rigoureux il faut cependant ajouter trois cas spéciaux. L'un apparaît en dimension 7, l'autre en dimension 8. Les objets que cela concerne ne sont pas classés mais forment un ensemble non vide, et enfin il y a le cas quaternionien. A ce jour, on ne connaît pas d'exemple non localement symétrique (ce sont les HP^n et Hyp_H^n). Cependant le cas à groupe d'holonomie G, caractérisé comme laissant invariante une 3-forme extérieure, semble prometteuse en physique mathématique (Witten).

6 Variétés riemannienne comme systèmes dynamiques

Faisons maintenant d'une VR un objet vivant. On poursuit les géodésiques indéfiniment, c'est-à-dire on se s'intéresse plus du tout aux géodésiques comme fournissant les plus courts chemins, mais comme les trajectoires de la mécanique hamiltonnienne. On demande ce qui va se passer quand le "temps", la longueur, tend vers l'infini. La dichotomie triviale est que, ou bien au bout d'un certain temps la géodésique revient au même point et avec le même vecteur vitesse, et alors elle est *périodique*. Sinon elle ne s'arrête jamais. Mais alors, il y a une multitude de cas possibles : elle peut s'accumuler autour d'une géodésique périodique, ou aller partout, mais alors est-ce de façon bien répartie, et il y a tous les cas intermédiaires. Ces problèmes sont ceux de la théorie ergodique, étudiée ici dans le cas particulier des VR. On reste dans le cas compact.

6.1 Géodésiques périodiques

Les géodésiques périodiques (GP) sont pour un physicien les états stationnaires, la longueur correspondant à l'énergie. Ce qui intéresse les physiciens est la fonction de comptage FC(L), désignant le nombre de GP qui sont de longueur inférieure ou égale à L. En général on espère seulement trouver une estimation asymptotique de FC(L) quand L tend vers l'infini.

Nous avons été bien naïfs de parler de FC(L), pire de son comportement asymptotique alors qu'aujourd'hui on ne sait même pas, à partir de la dimension 3, si une VR (quelconque) admet plus d'une géodésique périodique (on ne compte pas, dans cette vision seulement géométrique, les itérations de cette GP, on s'intéresse aux GP géométriquement distinctes).

En toute généralité, on sait seulement (1993) que pour les surfaces ($n = 2$) du type de la sphère ou du tore, il y a un nombre infini de GP et que FC(L) croît au moins comme $\frac{L}{\log L}$. Mais le comportement exact est inconnu pour toute surface générale qui n'est pas à courbure constante. Ici encore, les formes d'espaces à courbure négative sont les seules pour lesquelles on a des résultats significatifs. Pour toute surface ayant au moins deux trous, la fonction de comptage est exponentielle, et mieux, l'exposant de cette exponentielle est supérieur ou égale à celui que l'on rencontre pour les métriques à courbure constante, avec égalité seulement pour ces métriques (cette constante ne varie pas dans l'espace des modules) (1988). Ce résultat subsiste en dimension quelconque pour toutes les formes d'espaces à courbure négative (1991).

En dimensions $n > 3$, on ne sait établir l'existence que d'au moins une GP (1951). Pourquoi les choses sont-elles difficiles ? L'outil idéal devrait être la théorie de Morse : les GP sont les points critiques de la fonction longueur, fonction sur l'espace $\Omega(M^n)$ de toutes les courbes fermées de notre variété. Alors la théorie de Morse dit que l'on a au moins autant

de points critiques que les nombres de Betti de $\Omega(M^n)$. Or on sait (ce n'est pas facile et toujours en cours de raffinement) que ces nombres de Betti sont bien plus grands que ceux de M^n. Ce qui devrait nous combler, malheureusement il y a ici deux inconvénients majeurs à utiliser la théorie de Morse pour ce problème : d'abord il faut que les points critiques soient non dégénérés ; ensuite la théorie de Morse ne dit jamais, en toute généralité, quelles sont les longueurs correspondant à ces points critiques, et donc si ces longueurs sont toutes proportionnelles : on risque de tomber sur la même géodésique parcourue un nombre (infini) de fois. Peut-on se tirer d'affaire avec des hypothèses supplémentaires, mais raisonnables ?

Pour le découvrir, et en même temps sentir la difficulté du problème, deux exemples sont importants. Le premier est dû à Morse (1934), dans le cas tout simple des ellipsoïdes, valable en dimension quelconque. Pour tout $L > 0$, il existe un $\varepsilon(L)$ tel que pour tous les ellipsoïdes dont les axes (tous différents) diffèrent au plus de cet $\varepsilon(n)$, les GP, autres que celles d'intersection avec les plans de coordonnées, ont forcément une longueur supérieure à L. En quelque sorte avec des ε tendant vers zéro, toutes les GP s'évanouissent à l'infini, à part celles des plans de coordonnées

Le second exemple est dû à Alain Weinstein (1970). On considère les surfaces de révolution à courbure constante $+1$, elles ont deux points singuliers en leurs pôles. Selon que la longueur de l'équateur est dans $2\pi\mathbb{Q}$ ou non, toutes les géodésiques seront périodiques, ou aucune ne le sera. Il suffit d'adoucir autour des deux pôles, d'ajouter cela à une surface quelconque et on obtient une surface où la moindre variation change radicalement la nature des GP. Cet exemple marche en dimension quelconque.

D'où l'idée, dans l'ensemble de toutes les métriques d'une variété donnée, de considérer seulement des métriques génériques. Alors toutes les GP seront non dégénérées au regard de la théorie de Morse (1972). Mais il faut aussi contrôler les longueurs. Or depuis 1981 on sait que l'on a deux constantes positives , b telles que $FC(L) > \frac{a}{L} \sum_{k<bL} b_k(\Omega(M^n))$ (les b_k sont les nombres de Betti). Pour une métrique générique, on sait donc que $\mathrm{FC}(L)$ croît très vite. Enfin, les variétés compactes ne sont que de deux sortes, celles qui sont rationnellement elliptiques et celles qui sont rationnellement hyperboliques. Prises au hasard, elles ont toutes chances d'être hyperboliques ; on conjecture par ailleurs que les variétés à courbure positive sont elliptiques. Finalement on sait donc que, pour une métrique riemannienne générique sur une variété hyperbolique, la fonction $\mathrm{FC}(L)$ croît exponentiellement, et ce pour deux raisons : on sait que les nombres de Betti des $\Omega(M^n)$ croissent exponentiellement lorsque(M^n) est hyperbolique, plus le fait que la croissance exponentielle l'emporte sur la croissance linéaire des longueurs des GP itérées.

6.2 Le flot géodésique.

Le flot géodésique d'une VR est l'ensemble des $G_t : UM \prod UM$, où UM est le fibré unitaire tangent à notre VR, où t parcourt les réels, $G_t(v)$ étant le vecteur tangent en t à la géodésique dont le vecteur tangent en $t = 0$ est v. C'est un bon système dynamique, car le flot préserve la mesure de Liouville de UM.

"Que dire de l'application qui a une VR associe son flot géodésique parmi les flots sur les variétés ?". On sait qu'il n'y a pas toujours unicité en remarquant qu'il est utile de considérer des flots comme identiques (isomorphes) s'ils sont conjugués. Par contre, il y a unicité en

courbure négative (1990) ; de nombreux raffinements en cours, par exemple le spectre des longueurs (marqué par les classes d'homotopie) l peut même suffire à caractériser la variété. Dans le cas de courbure négative, le flot géodésique est toujours ergodique (et mieux, mixing) (1967).

Pour les géomètres qui trouvent les définitions classiques de l'entropie topologique *htop* difficiles à appréhender, on a depuis 1995 une très belle définition (ou propriété) de l'entropie du flot géodésique, à savoir que, pour toute VR (compacte) :

$$h_{top}(M^n; g) = \lim_{L\to\infty} \frac{1}{L} \int\limits_{M\times M} \log(FC(p,q;L))\, dpdq$$

où FC(p,q ;L) désigne le nombre de géodésiques de longueur inférieure à L qui joignent les points p et q.

A propos de la géométrie symplectique, qui devient un élément déterminant du paysage géométrique actuel et à venir, signalons qu'il ne semble pas que, dans le royaume symplectique, les flots géodésiques des VR soient vraiment spéciaux.

7 Le laplacien sur les fonctions : spectre, fonctions propres

On reste encore dans le cas compact. Une VR hérite automatiquement d'un invariant fondamental, le laplacien Δ opérant sur les fonctions numériques : $f \mapsto \Delta f$. C'est un opérateur elliptique du second ordre dont le symbole est la métrique elle-même, avec un signe négatif ici (mais la convention sur le signe varie avec les auteurs). Pour toute VR compacte, le laplacien fournit des invariants remarquables, les valeurs propres $\{\lambda_i\}$ (toutes positives, l'ensemble est appelé le spectre de la VR), et les fonctions propres associées (normalisées), ϕ_i, qui forment pour chaque λ_i un espace vectoriel de dimension finie, la multiplicité de cette valeur propre λ_i. Il est ainsi possible d'associer à toute VR une suite infinie de réels, et des fonctions numériques qui lui sont naturellement attachées (des informations géométriques comme leurs hypersurfaces nodales par exemple). On déduit classiquement de la théorie des opérateurs elliptiques, que l'on a une analyse de Fourier parfaite. Si l'on veut faire ici aussi de notre VR un objet vivant, on peut considérer l'équation de la chaleur, celle des ondes ou encore celle de Schrödinger associées à ce Δ. La connaissance du spectre et des fonctions propres permet, au moins théoriquement, de résoudre complètement ces trois équations.

A cause du rôle joué en physique par l'énergie, on s'intéresse en physique au comportement asymptotique du spectre, plus qu'à des valeurs exactes. Depuis 1953 on connaît le premier terme de ce développement, à savoir que

$$FC(\lambda) = \text{ nombre de } \{\lambda_i < \lambda\} \approx \bullet(n)\, Vol(g)\, \lambda^{n/2}.$$

Maintenant que dire de l'application "spectre" qui envoie une VR dans $\mathbb{R}^{\mathbb{Z}}$? Pour ce qui est de l'image, on sait (1987) que toute suite finie, peut être réalisée comme le début du spectre d'une VR convenable. Mais pour des suites infinies, on ne sait presque rien[10].

Pour ce qui est des images inverses (i.e. d'une unicité éventuelle) on a des résultats partiels. Il n'y a pas toujours unicité ; il existe même de nombreux exemples, y compris des

[10] voir cependant [28].

déformations à un paramètre de variétés isospectrales, dans tout voisinage de la métrique canonique sur la sphère : [12]. Par contre, il y a compacité de l'image inverse en dimension 2 mais seulement des résultats partiels en dimensions 3, 4 (1989-1994). Peu de choses sur le cas générique, où l'on serait peut-être en droit d'attendre plus, par exemple la finitude du nombre des paramètres.

Sur l'image on peut espérer obtenir plus dans l'esprit de la section III, par exemple sur la distribution du spectre dans $\mathbb{R}^+$, ou un contrôle plus précis. En fait la question essentielle est la répartition du spectre autour de la courbe continue que constitue l'asymptotique de Weyl en $\bullet\lambda^{n/2}$. On peut chercher des encadrements, ou ce que sont les sauts de la fonction en escalier $FC(\lambda)$. Depuis 1985 on a un double encadrement avec la seule courbure de Ricci, pour tout i :

$$\bullet(InfRicci, n, Diam)i^{2/n} < \lambda_i < \bullet(InfRicci, n, Vol)i^{2/n}$$

Ceci est possible parce qu'on a un contrôle isopérimétrique en fonction seulement de la borne inférieure de Inf Ricci. On a aussi un contrôle des trous en fonction du rayon d'injectivité pour les VR telles que $|K| < 1$. Cela semble raisonnable vu ce qui a été dit dans la section III. En fait, on ne le sait que depuis 1996 et seulement en dimension paire, et la démonstration est extrêmement compliquée.

Il existe très peu de travaux sur les fonctions propres elles-mêmes. On sait maintenant que leurs hypersurfaces de niveau sont toujours assez régulières (en un sens grossier certes), à savoir le comportement de leur volume : pour toute λ et toute fonction propre de cette valeur on a

$$c(g) \quad \sqrt{\lambda} < Vol\left(\phi^{-1}(0)\right) < c'(g)\sqrt{\lambda}$$

Il reste cependant encore des choses à faire : par exemple les constantes en question peuvent-elles être universelles ?

Nous avons regardé les implications sur le spectre d'hypothèses sur la courbure. Existe-t-il des relations entre le spectre du laplacien, ses fonctions propres, et le flot géodésique. Cette question intéresse beaucoup les physiciens qui l'appellent "l'approche semi-classique" : Peut-on récupérer la mécanique hamiltonnienne lorsque dans la mécanique quantique la constante de Planck dans l'équation de Schrödinger tend vers zéro ? Voici d'abord une réponse intéressante, mais assez théorique pour extraire des choses précises. On a une formule du type de celle de Poisson dans les $\mathbb{R}^n$, à savoir que pour une métrique générique :

$$\sum_i \exp\left(\pm\sqrt{-\lambda_i t}\right) = \sum_{L\in l} T_l$$

où les sommations portent sur et sont des distributions, les T_l étant des distributions attachées de façon précise aux GP. En sens inverse en quelque sorte, on sait attacher à certaines GP des suites infinies d'éléments du spectre, mais les choses sont-elles pleines dans certains cas ? (1977)

Comme on l'a souvent vu ci-dessus, on sait plein de choses remarquables dans le cas de la courbure négative et des formes d'espace. Une question un peu plus générale est de savoir si l'ergodicité du flot géodésique entraîne une excellente régularité pour les valeurs propres,

mais aussi pour les fonctions propres. Ceci n'est pas terminé, on ne sait pas, par exemple, si les mesures $\phi_i^2 dp$ convergent, lorsque i tend vers l'infini, vers la mesure totale dp. On ne le sait pour l'instant (1974, 1985) que pour une sous-suite (pleine certes) $i(k)$ du spectre.

Un cas très étudié est celui des formes d'espace, qui se révèle encore une fois un remarquable laboratoire de géométrie, parce qu'accessible par ordinateur. C'est le domaine du chaos quantique. On connait de nombreuses propriétés du spectre : pour les valeurs propres, on étudie si elles se distribuent comme les valeurs propres de matrices symétriques aléatoires (ensemble orthogonal gaussien GOE). On a des résultats de natures opposées, le fait que la forme d'espace soit arithmétique ou non (voir section II) jouent un rôle essentiel, quoique mal compris, et semble parfois paradoxal[11].

Actuellment, la question du scattering divise les experts. C'est une très belle question géométrique sur les rapports entre le flot géodésique et les fonctions propres : y a-t-il ou non une certaine accumulation des lignes nodales (dans le cas de dimension 2) autour des GP ?[12].

8 Quelles généralisations ?

Les VR sont des espaces excellents pour faire de la géométrie quoique suivent, comme on l'a vu à plusieurs reprises ci-dessus, certainement trop généreux. Par ailleurs, à d'autres égards, les VR ne sont pas assez générales. En voici quelques-uns, et la liste n'est pas du tout exhaustive.

De nombreux objets géométriques sont lisses presque partout, mais admettent des singularités, raisonnables et en nombre contrôlé ou encore les variétés d'orbites orbifolds, les variétés algébriques, les objets géométriques de la surface d'un polyèdre (polytope en dimension quelconque), la mécanique hamiltonnienne dans de nombreux cas ou encore ceux qui apparaissent dans certaines géométries sans rapport apparent avec la GR : les graphes (pour des métriques plus ou moins triviales), en particulier les groupes discrets (pour la métrique des mots), les immeubles de Tits.

Un autre besoin vient de la GR elle-même, à savoir : que se passe-t-il dans l'étude de compacité quand les quatre conditions de la section III.2 ne sont plus toutes les quatre remplies. Typiquement que se passe-t-il si on permet au volume de tendre vers **zéro,** à la courbure de tendre vers l'infini, du côté positif ou du côté négatif ?

Voici quelques-unes des réponses, des espaces inventés pour répondre à ces questions. Les premiers espaces sont ceux dit d'Alexandrov à courbure bornée inférieurement (par un k réel). Ce sont des espaces métriques, où l'on prend pour conditions de définitions des inégalités locales pour les triangles qui copient celles rencontrées pour les VR à K $>$ k. Mais pour avoir des espaces raisonnables il faut demander plus, typiquement dans les espaces à courbure positive ou nulle, que seulement la somme des angles des triangles soit toujours supérieure ou égale à π. Ne fût-ce que parce qu'il faut précisément disposer d'outils pour définir des angles ! Enfin il faut que la mesure de Hausdorff n soit finie. Ces espaces contiennent toutes les limites des VR à K $> k$. On contrôle parfaitement les singularités : leur ensemble est de dimension de Hausdorff inférieure ou égale à $n-1$, et le complément admet

[11] Voir les travaux de [36], [37], [29].
[12] voir les travaux de [36], [38].

une structure riemannienne C^0 (1994). Dans cette classe d'espaces on trouve en particulier, pour $k = 0$, la géométrie des frontières des polytopes des E^n. On peut aussi y développer toute une analyse [27]. Un livre très détaillé des fondements de la géométrie métrique est celui de D.[9].

On peut aussi chercher ce que sont les limites des VR avec cette fois-ci $\mathrm{K} < k$. On définit ce que l'on appelle des espaces CAT(k) (pour tout réel k). Encore ici on peut définir ces espaces métriques seulement *a priori* par des comparaisons entre leurs triangles et ceux de l'espace de référence $S^n(k)$. De même que ci-dessus dans le cas $k = 0$, il faut demander plus que la seule condition que la somme des angles des triangles soit toujours inférieure ou égale à π. On a besoin d'une définition qui permette de définir des angles. Le cas des CAT(0) est extrêmement important car il couvre en particulier les immeubles de Tits, les groupes discrets, les graphes (qui sont même des CAT(-∞))[13].

Une question très importante, en particulier pour la physique mathématique, est de compactifier la classe des métriques à Ricci > 0. La difficulté est d'obtenir à la fois une métrique et aussi une mesure, toutes les deux naturelles. Des travaux sont en cours par Kontsevitch.

Passons maintenant à deux situations géométriques extrêmes. La première est l'objet de la très longue section $3\frac{1}{2}$ de l'article de [15]. La notion de *mm*-espace est simple : il s'agit seulement d'espaces métriques (complets) munis en outre d'une mesure telle que la métrique soit une fonction mesurable. Dans ces espaces, les fonctions différentiables y sont remplacées par les fonctions lispschtziennes, et on peut définir des notions comme celles de courbure de spectre (en utilisant le principe minimax, puisque l'on a pas de laplacien). Ces espaces sont aussi de "bons" espaces métriques pour faire des probabilités.

L'autre situation extrême est d'étudier la géométrie des espaces avec d'autres données géométriques. Deux exemples (liés en partie) : celui des géométries symplectique et de contact, et celui des structures complexes. Pour plus de détails nous renvoyons à l'article de [17]. Beaucoup prophétisent que la géométrie du 21ième siècle s'appuiera beaucoup sur les géométries symplectique et de contact.

Même en restant à l'intérieur des *mm*-espaces, il y a un gouffre entre les plus généraux et la GR, voir même les espaces d'Alexandrov ou les CAT(k). Il va falloir le combler progressivement. Citons [14] :

"Il existe une large classe de problèmes provenant typiquement des sciences expérimentales (biologie, chimie, géophysique, médecine) où l'on a à traiter un nombre énorme de données, mais qui sont structurées de façon lâche. Les mathématiques traditionnelles, la théorie des probabilités et la mathématique statistique donnent des résultats satisfaisants quand le défaut de structure est grand. Paradoxalement, le manque d'organisation structurelle et de corrélations au niveau local conduit à un degré élevé de symétrie globale. Et par conséquent la loi de Gauss , qui permet de répartir les probabilités, surgit de la somme des variables aléatoires.

Mais nous rencontrons souvent des données structurées dans des situations où le calcul des probabilités ne s'applique pas. Ainsi les formations minéralogiques, ou encore des images de tissus vivants données par un microscope, abritent des corrélations inconnues qu'il faut cependant prendre en compte (car ce que nous croyons voir ne se révèlent jamais comme de

[13] Voir l'ouvrage entier qui leur est consacré : [8].

vraies images, mais le résultat de dispersion de certaines ondes : lumière, ultrasons, ondes sismiques....) Des exemples plus théoriques de structures de ce type apparaissent dans la théorie de la percolation qui rend compte des transitions de phase en physique, ou encore dans le mouvement brownien sans auto-intersection (qui modélise les longues chaînes moléculaires situées dans des solvants).

De tels problèmes, qui sont en quelque sorte déployés entre la symétrie pure et le chaos total, attendent l'émergence d'une sorte complètement nouvelle de mathématiques..."

Enfin nous sommes restés à un niveau toujours élémentaire de la géométrie. Il existe de plus en plus de géométries, en outre indispensables, dans lesquelles les notions d'espaces, de points, interviennent à un degré supérieur : un point sera toujours défini en même temps que tout un ensemble d'objets mathématiques qu'il doit transporter avec lui[14].

[14] voir [10].

Bibliographie

[1] *Discussion on Geometry*, dans Connes, Gromov, Milman (Eds.), 2000 Visions in Mathematics towards 2000, 895-915, Birkhäuser.

[2] M. Berger, *Riemannian geometry during the second half of the century*,(University Lecture Series, volume 17, second printing), AMS, 2000.

[3] M. Berger, *Riemannian geometry during the second half of the century* (updated and indexed version of Berger 1998a), Americ. Math. Soc, Providence, 2000.

[4] M. Berger, *Peut-on définir la géométrie aujourd'hui ?*, result. Math. 40, pages 37-87, 2001.

[5] M. Berger, *A panoramic view of Rienmannin geometry*, Springer, 2003.

[6] Besse, *A. Einstein manifolds*, Springer, 1987.

[7] A. Borel, *Compact Clifford-Klein forms of symmetric spaces*, Topology, 2, 111-122, 1963.

[8] M. Bridson & Haefliger, *A. Metric spaces of non-positive curvature*, Springer, Heidelberg, 1998.

[9] D. Burago, Y. Burago & S. Ivanov, *A course in Metric Geometry*, American math. Soc, 2001.

[10] P. Cartier, *A mad day's work : From Grothendieck to Connes et Kontsevich The evolution of concepts of space and symmetry*, Bulletin of the AMS, 38, 389-408, 2001.

[11] S. Donaldson & P. Kronheimer, *The geometry of four-manifolds*, Oxford University Press, 1996.

[12] C. Gordon, *Isospectral deformations of metrics on spheres*, Inventiones math., 145, 317-332, 2001.

[13] M. Gromov, *Carnot-Caratheodory spaces seen from within*, in A. Bellaïche & J.-J. Risler (Eds.), Sub-Riemannian geometry Birkhaüser, 1994.

[14] M. Gromov, *Possible trends in mathematics in the coming decades*, Notices of the American math. Soc., 45, 846-847, 1998.

[15] M. Gromov, *Metric structures for Riemannian and non Riemannian manifolds*, edited by J. Lafontaine et P. Pansu. Birkhäuser, 1999.

[16] M. Gromov, *Local and global in geometry*, 1999.

[17] M. Gromov, *Spaces and questions*, in Connes, Gromov, Milman (Eds.), Visions in mathematics, GAFA 2000, towards 2000 , 118-161, Birkhaüser, 2000.

[18] M. Gromov, *Mesoscopic curvature and hyperbolicity*, No. IHES/M/01/26). IHES, 2001.

[19] M. Gromov, *Cat(k)-spaces : construction and concentration*, No. IHES/M/01/24). IHES, 2001.

[20] M. Gromov & Piatetski-Shapiro, *I. Non-arithmetic groups in Lobachevsky spaces*, Publ Math. Inst. Hautes Etud. Sci., 66, 93-103, 1988.

[21] M. Gromov & R. Schoen, *Harmonic maps into singular spaces and p-adic superrigidity for lattices in groups of rank one*, Publ. Math. IHES, 76, 165-246, 1992.

[22] E. Hebey, *Introduction à l'analyse non linéaire sur les variétés*, Diderot, 1997.

[23] N. Hitchin, *Global differential geometry*, in Mathematics unlimited : 2001 and beyond B. Engquist and W. Schmidt eds. 577-591. Springer, Heidelberg-New York, 2001.

[24] H. Hopf, *Differential Geometrie und topologische Gestalt*, Jahrbericht Math. Ver,DMV, 41, 209-229, 1932.

[25] G. Huisken, *Evolution equations in geometry*, in S. Engquist (Eds.), Mathematics Unlimited : 2001 and beyond (pp. 593-604). Springer, 2001.

[26] M. Kapovich, *Hyperbolic manifolds and discrete groups*, Birkhaüser, 2001.

[27] K. Kuwae, Y. Machigashira & T. Shioya, *Beginning of analysis on Alexandrov spaces*, Contemp. math., 258, 275-284, 2000.

[28] J. Lohkamp, *Discontinuity of geometric expansions*, Commentari. math. helv., 71, 213-228, 1996.

[29] W. Luo & P. Sarnak, *Quantum ergodicity of eigenfunctions on* $PSL2(Z)\mathbb{H}$, Publications Math. I. H. E. S., 81, 207-237, 1995.

[30] G. Margulis, *Discrete groups of motions of manifolds of non-positive curvature*, A M S Translations, 39, 33-45, 1977.

[31] G. D. Mostow, *A remarkable class of polyhedra in complex hyperbolic space*, Pacific J. of Math., 86, 171-276, 1980.

[32] A. Nabutovsky, *Einstein structures : existence versus uniqueness*, Geom Anal. functional Anal 5, 76-91, 1995.

[33] A. Nabutovsky & S. Weinberger, *Variational problems for Riemannian functionals and arithmetic groups*, Publ Math. IHES, 92, 5-62, 2000.

[34] P. Petersen, *Aspects of global Riemannian geometry*, Bulletin of the AMS, 36, 296-344, 1999.

[35] M. Postnikov, *Riemannian Geometry*, Springer. Heidelberg, 2001.

[36] Z. Rudnick & P. Sarnak, *Zeros of princial zeta-functions and random matrix theory*, Duke Math. J., 81, 269-322, 1996.

[37] P. Sarnak, *Arithmetic Quantum Chaos*, In Isr. math. conf. (183-236), 1995.

[38] Y. Shimizu & A. Shudo, *Polygonal billiards : corespondence between classical trajectories and quantum eigenstates*, Chaos, solitons and fractals, 5, 1337-1362, 1995.

[39] J. Tits, *Sur certaines classes d'espaces homogènes de groupes de Lie*, Mémoires Académie des Sciences de Belgique, XXIX, 1955.

[40] E. van den Ban, M. Flensted-Jensen & H. Schlichtkrull, *Harmonic analysis on semi-simple symmetric spaces : a survey of some genreal results*, In Pro symposia pure math... AMS **? ? ?.** (191-217), 1997.

[41] J. Wolf, *Spaces of constant curvature*, (second ed.), J. Wolf Berkeley, 1972.

La géométrie différentielle d'Élie Cartan à Charles Ehresmann et André Lichnerowicz

Paulette Libermann

1 Introduction

Élie Cartan (1869-1951) fut l'un des plus éminents mathématiciens de son époque. Son œuvre présente une grande unité ; elle gravite principalement autour de la théorie des groupes de Lie et de la Géométrie différentielle (locale ou globale) : systèmes différentiels, groupes finis et "continus", "groupes infinis", espaces généralisés, Géométrie riemannienne et pseudo riemannienne etc. ainsi que l'application de ces théories à la Mécanique analytique et à la Relativité. Grâce à une intuition géométrique remarquable et une grande aptitude à dominer les calculs les plus compliqués, il traita de façon exhaustive de nombreux exemples à partir desquels il a développé ses idées, plus préoccupé d'exposer des résultats nouveaux que d'approfondir les fondements de ses théories. Ses idées étaient en avance sur celles de son époque et le langage des années 1920-1930 n'était pas adapté à leur énoncé. Ses idées ont eu un tel impact sur le développement ultérieur de la Géométrie différentielle que nous aborderons des sujets antérieurs à 1930.

A l'exception d'Henri Poincaré qui écrivit peu avant sa mort un rapport sur les travaux d'Élie Cartan à l'occasion de la candidature de celui-ci à la Sorbonne, les mathématiciens français ne voyaient pas l'importance de l'œuvre. C'est à l'étranger et particulièrement en Allemagne, sous l'influence de Hermann Weyl , que l'on a reconnu, à partir de 1925, la valeur des travaux d'Élie Cartan . En 1929, lors d'un séjour en France, Albert Einstein rencontra Élie Cartan . Il s'en suivit, de 1929 à 1932, une longue correspondance sur le parallélisme absolu (voir [53]).

Les travaux d'Élie Cartan concernant les invariants intégraux et l'*Analysis situs* (ce qu'on appelle maintenant Topologie) furent connus, en dehors des publications de leur auteur, par la thèse de G de Rham (parue en 1951) et celle de C. Ehresmann publiée aux Etats-Unis en 1934.

En France le groupe Bourbaki organise en 1936 une série d'exposés sur les travaux d'Élie Cartan. Rédigés en tenant compte des progrès de l'Algèbre moderne et de la Topologie,

ces exposés firent beaucoup pour la diffusion des idées d'Élie Cartan à travers le monde, notamment en Europe de l'Est et aux États-Unis.

Parmi les idées nouvelles apportées par Élie Cartan , notons l'introduction de l'algèbre des formes différentielles extérieures sur une variété, accompagnée de la différentiation d. Cet opérateur d, qui vérifie la relation $dd = 0$ et augmente le degré d'une unité, est utilisé localement pour étudier notamment le problème d'équivalence locale et globalement pour définir les groupes de cohomologie. Avant É. Cartan les mathématiciens traitant d'*Analysis situs* (B. Riemann, H. Poincaré) ont surtout considéré l'homologie. Sous l'influence des travaux d'É. Cartan (notamment "Sur les invariants intégraux de certains espaces homogènes"), G. de Rham a démontré l'isomorphisme entre les groupes d'homologie et groupes de cohomologie ; ce sont les fameux théorèmes dits de "de Rham".

Le présent article est divisé en 3 parties. Dans la *premi*ère nous traiterons des travaux *d'Élie Cartan* [11] en nous limitant à un nombre restreint de sujets (systèmes différentiels, invariants intégraux, "groupes infinis", méthode du repère mobile, "espaces généralisés", ces sujets ayant un impact sur les travaux de C. Ehresmann et A Lichnerowicz). Pour un exposé plus complet sur l'œuvre d'É. Cartan, voir l'article consacré à celui-ci dans "Travaux Mathématiques" fascicule VIII, 1996, Publications du centre Universitaire du Luxembourg. Cet article contient une longue bibliographie. Voir aussi dans [11] l'introduction d'É. Cartan.

La *deuxième partie* traitera des travaux de *C. Ehresmann* (espaces fibrés, connexions, pseudogroupes, jets, groupoïdes de Lie). La théorie des feuilletages ne sera pas traitée ; C. Ehresmann et G. Reeb y ont fait un travail de pionniers, personne avant eux n'ayant abordé ce sujet ; cette théorie qui n'a pas été directement inspirée par É. Cartan est développée notamment par A. Haefliger dans [24].

La *troisième partie* n'abordera qu'un des aspects des travaux *d'A. Lichnerowicz*. L'œuvre si vaste de celui-ci pourrait faire l'objet d'autres articles. Nous exposerons seulement sa théorie des *variétés de Poisson et de Jacobi*, théorie actuellement en plein développement.

2 Première partie : sur les travaux d'Élie Cartan (1869-1951)

2.1 Systèmes différentiels et calcul différentiel extérieur

Pendant la période 1899-1914, Élie Cartan s'est notamment consacré à l'étude des systèmes différentiels et des "groupes infinis" en introduisant le calcul différentiel extérieur, généralisant la notion de "covariant bilinéaire" due à Frobenius et à Darboux.

A la forme de Pfaff (ou forme de degré 1) $\phi = \sum_{i=1}^{n} A_i dx^i$, il associe la forme $d\phi = \sum_{i,j} \frac{\partial A_i}{\partial x^i}\, dx^j \wedge dx^i$. Cela lui permet d'étudier dans $\mathbb{R}^n$ (ou $\mathbb{C}^n$) les équations de Pfaff (chaque équation étant définie par $\phi = 0$ où ϕ et une forme de Pfaff) et les systèmes de Pfaff $\phi^1 = 0, \ldots, \phi^n = 0$ (où $\phi^1, \ldots, \phi^n$ sont des formes de Pfaff linéairement indépendantes). En utilisant la différentielle extérieure de ces formes il démontre que, moyennant des conditions de régularité, une équation de Pfaff est de *classe impaire* $2n+1$ c'est-à-dire s'exprime comme :

$$dz + \sum_{i=1}^{n} y^i dx^i = 0$$

Élie Cartan développe sa théorie de l'*involution* (relative aux systèmes de Pfaff). Un système de Pfaff Σ est dit *involutif* s'il existe en tout point x une chaîne $E_1 \subset E_2 \subset \ldots \subset E_n$ d'éléments de contact intégraux de Σ. Un élément de contact est le couple (x,π) d'un point x de $\mathbb{R}^n$ (ou $\mathbb{R}^n$) et d'une variété linéaire affine π de $\mathbb{R}^n$ (ou $\mathbb{C}^n$) passant par x. L'élément de contact (x,π) est intégral de la forme $\phi = \sum_{j=1}^{n} A_j dx^i$ si la restriction de ϕ à tout vecteur appartenant à (x,π) est nulle. Élie Cartan a démontré un théorème d'existence (de caractère local) de variétés intégrales d'un système de Pfaff Σ en involution ; une variété intégrale de Σ, de dimension q, est l'image d'une application analytique $f = (f^1, \ldots, f^n)$ d'un ouvert de $\mathbb{R}^q$ (ou $\mathbb{C}^q$) dans R^n (ou $\mathbb{C}^n$) telle que la restriction de Σ à cette image soit nulle. L'involution permet de démontrer des théorèmes d'existence d'équations aux dérivées partielles. Par exemple une équation aux dérivées partielles dans $\mathbb{R}^3$ s'écrivant

$$F(x,y,z,p,q) = 0 \quad \text{avec} \quad p = \frac{\partial z}{\partial x} \quad \text{et} \quad q = \frac{\partial z}{\partial y}$$

peut être considérée comme le système différentiel dans $\mathbb{R}^5$ (de coordonnées x, y, z, p, q)

$$\begin{cases} F & = & 0 \\ dz - pdx - qdy & = & 0. \end{cases}$$

C'est grâce à sa théorie des systèmes en involution qu'Élie Cartan a pu démontrer la compatibilité des systèmes d'équations aux dérivées partielles introduites en Relativité par Albert Einstein.

Nous en arrivons à un problème qui a intéressé Élie Cartan dans une grande partie de son œuvre et qu'il a appelé le *problème d'équivalence* ; dans le cas présent le problème d'équivalence des systèmes de Pfaff est la recherche des conditions pour que deux tels systèmes se ramènent l'un à l'autre par changement de variables, ce qui conduit à attacher à chaque système de Pfaff un ensemble d'invariants (fonctions ou constantes) qui le caractérise ; par exemple la classe pour une équation de Pfaff.

A la suite des travaux de Kähler, Élie Cartan a étendu sa théorie de l'involution à des systèmes différentiels de différentiels de degré quelconque [10].

2.2 Groupes continus "infinis"

Les questions étudiées dans le paragraphe précédent ont leurs applications dans la théorie des "groupes infinis continus" initiée par S. Lie, poursuivie par Engel, Medolaghi et Vessiot ; É. Cartan développe cette théorie de 1904 à 1909 dans quatre gros articles parus aux Annales de l'École Normale et qu'il résume dans les exposés au Séminaire Julia en 1937.

Un "groupe infini continu" est un ensemble Γ d'homéomorphismes analytiques d'ouverts de $\mathbb{C}^n$(ou $\mathbb{R}^n$), muni d'une loi de composition partiellement définie, Γ constituant la solution générale d'un système S d'équations aux dérivées partielles analytique, assujetti à certaines conditions de régularité (dont l'ordre est appelé l'ordre du "groupe infini") ; si S est un système de Mayer-Lie, c'est-à-dire si toutes les dérivées partielles d'ordre $\geq r$ s'expriment en fonction des dérivées d'ordre inférieur, la solution dépend d'un nombre fini de paramètres et l'on retrouve un "groupe fini continu". Alors que ses prédécesseurs ont essayé d'aborder la théorie au moyen de transformations infinitésimales, É. Cartan développe un point de vue

différent pour définir les notions d'isomorphisme, "sous-groupe infini invariant", "groupe infini simple" ; la difficulté est qu'en général Γ' n'admet pas de structure de groupe. É. Cartan introduit la notion de *prolongement* (technique utilisée souvent dans son œuvre et qui consiste à introduire de nouvelles variables) : Γ' opérant localement dans $\mathbb{C}^{n+p}$(resp. $\mathbb{R}^{n+p}$) est le prolongement de Γ opérant localement dans $\mathbb{C}^n$ (resp. $\mathbb{R}^n$), si tout $f' \in \Gamma'$ se projette suivant une application $f \in \Gamma$; le prolongement est holoédrique, si à tout $f \in \Gamma$ correspond un seul $f' \in \Gamma'$; il est mériédrique dans le cas contraire.

É. Cartan a obtenu les trois théorèmes fondamentaux suivants, généralisant ceux de Lie pour les groupes finis.

1. Tout "groupe infini" admet un prolongement holoédrique Γ' du premier ordre (prolongement normal) opérant dans $\mathbb{C}^{n+p}$ (ou $\mathbb{R}^{n+p}$) et caractérisé par la propriété de laisser invariantes des formes de Pfaff dans $\mathbb{C}^n$(ou $\mathbb{R}^n$).
2. Ces formes ω^i vérifient les équations de structure

$$d\omega^i = \sum_{j,k} c^i_{jk}\omega^j \wedge \omega^k \sum_{j,\alpha} a^i_{j\alpha}\omega^j \wedge \bar{\omega}^\alpha, \tag{1.4}$$

   les formes $\bar{\omega}^\alpha$ étant définies dans $\mathbb{C}^{n+p}$ (ou $\mathbb{R}^{n+p}$), les coefficients $a^i_{j\alpha}$ (nuls dans le cas fini) définissant un système involutif.
3. A la donné des coefficients $c^i_{j\alpha}$ et $a^i_{j\alpha}$ vérifient certaines conditions de compatibilité, il correspond un "groupe infini".

É. Cartan a déterminé les "groupes infinis" isomorphes holoédriques d'un "groupe infini" donné et défini un "groupe infini simple" comme un groupe infini n'admettant pas de "groupe infini" isomorphe mériédrique. Il a obtenu la classification, en quatre classes, des "groupes infinis" simples transitifs complexes.

Depuis 1952, avec l'introduction des pseudogroupes de Lie et des jets par C. Ehresmann, ces questions ont pris un développement considérable. La théorie de Cartan a été formulée sous forme globale (un "groupe infini" est un noyau de pseudogroupe de Lie).

2.3 Géométrie Différentielle et ses applications à la Mécanique Analytique et à la Relativité

1. Reproduisant un cours professé à la Sorbonne en 1920-1921 dans le cadre d'un enseignement de Mécanique analytique et Mécanique céleste, Élie Cartan [7] a écrit un livre *Leçons sur les invariants intégraux* publié en 1922 et réédité plusieurs fois. Nous montrerons dans la suite (voir les structures de Poisson) l'importance de ce livre dont le style paraît moins "daté" que celui d'autres ouvrages publiés ultérieurement. La théorie des invariants intégraux a été fondée par H. Poincaré et exposée par lui dans le tome III de ses *Méthodes nouvelles de la Mécanique Céleste*. L'apport nouveau d'É. Cartan a été l'utilisation de l'algèbre extérieure et l'adjonction du temps à l'invariant intégral de Poincaré. Ainsi se plaçant dans l'espace-temps, produit de l'espace $\mathbb{R}^n$ par R (représentant le temps), É. Cartan introduit à la base de la Mécanique le principe de la conservation de la quantité de mouvement et de l'énergie, exprimé au moyen d'une forme de Pfaff ω ; cette forme ω, appelée par l'auteur "tenseur de quantité de mouvement – énergie" et désignée actuellement par "invariant intégral de Cartan" est

définie dans l'espace-temps ; elle s'écrit :

$$\omega = \sum_{i=1}^{n} p_i dq^i - H dt,$$

où les q^i sont des coordonnées dans $\mathbb{R}^n$, t la variable de temps, les p_i sont les quantités de mouvement du système et H son énergie. Dans cette formulation, le temps ne joue plus un rôle privilégié ; les conditions obtenues sont indépendantes des changements de coordonnées dans l'espace-temps.
É. Cartan utilise sa théorie pour traiter du problème des n corps en Mécanique céleste et des principes de l'optique.

2. Parmi les contributions d'É. Cartan à l'étude des groupes de Lie, figure sa *théorie du repère mobile*.
Soit G un groupe de Lie, T_eG l'espace tangent en $e \in G$; à une base $(\epsilon_1, \ldots, \epsilon_p)$ de T_eG correspondent p champs de vecteurs invariants à gauche définissant en chaque point s de G une base de T_sG. On obtient ainsi un exemple de repère mobile. L'originalité d'É. Cartan a été de considérer la base duale de $(\epsilon_1, \ldots, \epsilon_p)$, définissant ainsi des formes *globales invariantes à gauche* sur G. L'auteur a montré que ces formes vérifient les équations (appelées maintenant équations de Maurer-Cartan)

$$d\omega^\alpha = \sum c^\alpha_{\beta\gamma} \omega^\beta \wedge \omega^\alpha \quad \alpha = 1, \ldots, p$$

où les $c^\alpha_{\beta\gamma}$ sont les constantes de structure du groupe G.
É. Cartan a ensuite étudié les *espaces homogènes* de Klein c'est-à-dire les variétés du type $M = G/H$, où H est un sous-groupe fermé de G. Sa théorie du repère mobile s'est montrée très efficace, notamment dans l'étude des sous-variétés d'un espace homogène. Il a traité de ce sujet dans de nombreux articles et dans son livre *La théorie des groupes finis et continus et la géométrie différentielle traitées par la méthode du repère mobile* d'après la rédaction de G. Leray [8].
É. Cartan choisit d'abord un repère "fixe" ou repère référence R_0, qui peut être par exemple une famille finie de points de l'espace homogène $M = G/H$; il considère l'ensemble $(R_s)_{s\in G}$ des transformés de R_0 par les éléments de G, la condition suivante devant être remplie : pour tous $s \in G$, $u \in G$, il existe une seule transformation de G envoyant R_s sur R_u.
L'action à gauche $(s, x) \to sx$ de G sur M induit pour tout $x \in M$ une application

$$\Phi_x : G \times G \to M$$

telle que

$$\Phi_x(s, u) = s^{-1}ux;$$

on a $\Phi_x(s, s) = x$. Soit $\Phi_{x,s}$ la restriction de Φ_x à $\{s\} \times G$. Alors l'application linéaire tangente $T_s\Phi_{x,s} : T_sG \to T_xM$ est ce que É. Cartan appelle "la transformation analytique infinitésimale de R_s dans R_{s+ds} rapporté au repère R_s".
Pour tout élément X de l'algèbre de Lie $\mathcal{G}$ de G, cette application envoie $-X^R(s)$ sur $X_M(x)$, où X^R et X_M sont respectivement le champ invariant à droite sur G et le

champ de vecteurs fondamental sur M associés à X. On obtient ainsi une forme sur G à valeurs dans l'algèbre de Lie des champs de vecteurs fondamentaux. Voir [5] et [42]. Par exemple si $G =$ groupe des déplacements euclidiens de E^n et $H = SO(n)$, le repère de référence dans E^n est formé d'un point et d'une base orthonormée. Un déplacement euclidien amène $(x_0, (e_1)_0, (e_2)_0, \ldots, (e_n)_0)$ sur $(x, e_1, e_2, \ldots, e_n)$. On écrit :

$$dx = \sum_i^n \omega^i e_i$$

$$de_i = \sum_i^n \omega_i^j e_j,$$

où les formes ω_i^j sont les composantes de la "transformation infinitésimale rapportée au repère mobile". Par rapport au repère "fixe", on a :

$$de_i = \sum_i^n \omega_i^j s_j^k (e_k)_0$$

3. Depuis 1870, la Géométrie était en partie dominée par les idées du programme d'Erlangen énoncé par F. Klein concernant les espaces homogènes. Mais les continuateurs de l'œuvre de Riemann contribuaient aussi au développement de la Géométrie. La découverte en 1916 de la Relativité générale par A. Einstein, le développement du calcul différentiel absolu (c'est-à-dire les lois de dérivation covariante) à la suite notamment des travaux de Ricci sur le calcul tensoriel et de Levi-Civita sur le déplacement par parallélisme, donnèrent une impulsion nouvelle à la Géométrie.
   Le désir de faire une synthèse entre ces deux points de vue a conduit Élie Cartan à introduire les "*espaces généralisés*" qu'il appelle d'abord "espaces de Klein déformés" et dont il donne une définition empirique. Un tel espace est "un espace qui, au voisinage de chacun de ses points, présente tous les caractères d'un espace de Klein et qui est doué de plus d'une loi permettant de raccorder en un seul espace les deux petits morceaux qui entourent deux points infiniment voisins". Pour expliciter ses idées, il est amené à préciser la notion de variété.
   Pour ces espaces, l'auteur définit la notion de *connexion* (linéaire, affine, conforme, projective) ; pour les connexions de Levi-Civita sur une variété riemannienne il introduit le *groupe d'holonomie* obtenu en relevant un chemin fermé de la variété dans l'espace des repères. Cette notion de groupe d'holonomie prendra une grande importance dans l'étude des espaces fibrés et feuilletés.
4. Elie Cartan a abordé l'étude de la *Relativité* dès 1922 dans un article "Sur les équations de la gravitation d'Einstein" (*J. Math pures et appliquées* t.1, p.141-203) et dans des Notes aux comptes rendus. Il introduit la notion d'espace riemannien à *parallélisme absolu*, notion retrouvée en 1928 par A. Einstein et sur laquelle celui-ci a cherché à fonder une théorie unitaire du champ gravitationnel - électromagnétique.
   La théorie de ces espaces à parallélisme absolu (c'est-à-dire des variétés M telles que le fibré tangent TM s'identifie à $M \times \mathbb{R}^n$) a été reprise dans ses articles "Sur les variétés à connexion affine et la théorie de la Relativité généralisée" (*Ann. École. Normale* 1923

et 1924). Pour ces espaces le déplacement par parallélisme d'un vecteur entre 2 points est indépendant du chemin choisi, d'où une courbure nulle.
C'est dans ces articles qu'Élie Cartan utilise sa théorie des systèmes différentiels extérieurs et de l'involution pour démontrer la compatibilité des systèmes d'équations aux dérivées partielles introduits par A. Einstein.
De 1929 à 1932, A. Einstein et É. Cartan échangèrent une longue correspondance sur le parallélisme absolu ("Letters on absolute parallelism") où Élie Cartan répond avec beaucoup de pédagogie aux interrogations de son correspondant [9] (voir sur ce sujet [53]).

3 Deuxième partie : Sur les travaux de C. Ehresmann (1905-1979)

Ces travaux, ainsi que ceux relatifs à la Topologie algébrique, sont publiés dans le volume I (parties 1,1 et 1,2) de ses œuvres complètes (voir Ehresmann [19]). Ce volume contient en outre une notice de l'auteur sur ses travaux rédigée en 1955 ainsi que des commentaires résumant et mettant en valeur l'essentiel de l'œuvre de
C. Ehresmann. Ces commentaires, en ce qui concerne la Géométrie différentielle et la Topologie algébrique sont dûs à W.T. Van Est (espaces homogènes et groupes de Lie), M. Zisman (espaces fibrés), G. Reeb (variétés feuilletées), P. Libermann (connexions et pseudogroupes de Lie), R. Thom (théorie des jets et développements ultérieurs), J. Pradines (groupoïdes différentiables), R. Hermann (applications à la Physique et à la théorie du contrôle). Enfin un article de A. Haefliger expose l'ensemble de l'œuvre. C'est pourquoi nous nous contenterons d'exposer brièvement certains des travaux de C. Ehresmann et renvoyons le lecteur aux textes ci-dessus (voir également [6]).

Une des motivations de C. Ehresmann pour élaborer les fondements de la géométrie différentielle fut de comprendre du point de vue global les idées d'Élie Cartan.

3.1 Espaces fibrés

L'auteur fut l'un des fondateurs de cette théorie ; beaucoup de notions qui paraissent "classiques" lui sont dûes. Ses recherches sur les espaces homogènes puis le désir de définir globalement les connexions et le transport par parallélisme le conduisirent à étudier les espaces fibrés.

Ses premiers travaux sont exposés dans une série de Notes aux Comptes-rendus publiées de 1941 à 1943, la première en collaboration avec J. Feldbau (où il démontre la théorie du relèvement des homotopies). Il précise la définition de *fibré pricipal*, *fibré associé* et montre l'importance du *groupe structural* G. Un des problèmes importants est la *réduction* du groupe structural $\hat{G}$ d'un fibré principal à un sous-groupe fermé G. L'auteur montre que ce problème se ramène à l'existence d'une section dans un fibré associé de fibre type $\hat{G}/G$, ce qui implique des conditions topologiques. Ainsi le groupe structural du fibré des repères $H(M)$ d'une variété M, de dimension n, est le groupe linéaire $GL(n,\mathbb{R})$. C. Ehresmann a étudié la réduction de $GL(n,\mathbb{R})$ à un sous-groupe G. Toute variété paracompacte M admet une structure *Riemannienne* car l'espace $GL(n,\mathbb{R})/O(n)$ étant contractile, le groupe

$GL(n,\mathbb{R})$ peut se réduire à $O(n)$. En revanche il n'existe pas nécessairement de métrique de type (p,q). L'auteur démontre (Comptes-rendus, 1943, p.628) qu'une variété compacte de dimension 4 ne peut admettre de forme quadratique de type $(x^1)^2-(x^2)^2-(x^5)-(x^4)$ que si sa caractéristique d'Euler-Poincaré est nulle. *Ainsi un univers possible en Relativité ne peut être* S^4 *ou* $P^4(\mathbb{R})$. Dans cette même Note, C. Ehresmann définit une structure de variété différentiable à l'aide d'un atlas de cartes locales, ce qui est devenu classique.

Une G-réduction du groupe $GL(n,\mathbb{R})$ a été appelée "structure infinitésimale régulière" par C. Ehresmann puis plus tard G-structure par S. Chern [7]. Voir aussi [4][54].

C'est le problème de la recherche des variétés différentiables admettant une structure complexe subordonnée qui a conduit C. Ehresmann à introduire *les structures presque complexes* définies sur une variété M_{2n} par un champ d'endomorphismes I du fibré tangent tels que $I^2=-id$.

C. Ehresmann a montré que l'existence d'une structure presque *complexe* est équivalente à celle d'une *structure presque symplectique* ; une telle structure est définie par une 2-forme différentielle Ω de rang $2n$. Si $d\Omega=0$, on a les structures symplectiques dont l'étude a connu un grand développement en raison de son importance en Mécanique. Dans ses exposés au Colloque de Topologie Algébrique de Paris (1947) et au Congrés International de Harvard (1950) il développe sa théorie et étudie les conditions topologiques d'existence d'une structure presque complexe ; par exemple pour une variété orientable de dimension 6, une condition nécessaire et suffisante est l'annulation de la troisième classe de Stiefel-Whitney ; ainsi la sphère S^6 possède une structure presque complexe (que l'on peut définir au moyen des octaves de Cayley) alors qu'on démontre que S^4 n'en possède pas. Ultérieurement il a été démontré par A. Borel et J. P. Serre que seules S_2 et S_6 admettaient des structures presque complexes.

3.2 Connexions infinitésimales et connexions "de Cartan"

Dans son célèbre article sur les connexions infinitésimales (colloque de Topologie de Bruxelles, 1950), fondamental dans l'étude ultérieure des connexions, C. Ehresmann définit une connexion dans un espace fibré par la donnée d'un champ d'éléments de contact transverse aux fibres, assujetti à certaines conditions. Une connexion principale sur un G-fibré principal P est définie par un tel champ qui doit être invariant par les translations à droite de G. L'auteur montre que ceci est équivalent à la donnée d'une *forme de connexion* ω sur P, à valeurs dans l'algèbre de Lie $\mathcal{G}$ de G, satisfaisant les conditions suivantes, exprimées avec les notations d'Élie Cartan par :

$$\begin{aligned} \omega(y+dy)s &= s^{-1}(y+dy)s \\ \omega(y(s+ds)) &= s^{-1}(s+ds) \end{aligned} \tag{1.5}$$

où $y+dy$ (resp. $s+ds$) est un vecteur tangent à P (resp. G) en y (resp. s) ; la même notation désigne à la fois la translation $y\to ys$ et son prolongement aux vecteurs.

C. Ehresmann définit la courbure (obstacle à la complète intégrabilité du champ transverse), la différentiation covariante ainsi que le groupe d'holonomie (en relevant les chemins fermés de la base). Il montre que si la courbure est nulle la connexion détermine un homomorphisme du groupe de Poincaré de la base B sur le groupe d'holonomie définissant chaque variété intégrale du champ horizontal comme un revêtement de B.

C. Ehresmann donne une définition des "espaces généralisés" d'Élie Cartan et des *connexions de Cartan*.

Un *"espace généralisé"* est un espace fibré différentiable E, de base B, de fibre type F, groupe structural $\hat{G}$, de fibré principal associé P possédant les propriétés suivantes :

1^0) La fibre type F est un espace homogène $\hat{G}/G$ et $dim\, F = dim\, B$.

2^0) Il existe une *soudure* de l'espace fibré à sa base au sens suivant : il existe une section σ de $\pi : E \to B$ induisant pour tout $x \in B$, un isomorphisme de l'espace tangent T_xB sur l'espace tangent $T_{\sigma(x)}E_x$, où $E_x = \pi^{-1}(x)$.

Les données précédentes entraînent l'existence d'un sous-fibré principal P de $\hat{P}$, de groupe structural $\hat{G}$. Ce fibré P est l'ensemble des difféomorphismes de $\hat{G}/G$ sur les fibres de F envoyant le point fixe de F sur les points de $\sigma(B)$.

Le fibré principal P est muni d'une *forme de connexion de Cartan* c'est-à-dire d'une forme $\bar{\omega}$ sur P à valeurs dans $\mathcal{G}$ (algèbre de Lie de $\hat{G}$), dont la restriction à T_yP (pour tout $y \in P$) est un isomorphisme sur $\mathcal{G}$, cette forme vérifiant des conditions analogues à (1). On en déduit un parallélisme sur P.

Remarquons que la forme $\bar{\omega}$ n'est pas une forme de connexion car elle est à valeurs dans $\hat{\mathcal{G}}$ mais elle s'étend en une connexion principale sur $\hat{P}$.

Les espaces à connexion affine, projective et conforme introduits par É. Cartan entrent dans ce cadre.

Par exemple l'espace des repères $H(M)$ d'une variété M a pour groupe structural $GL(n, \mathbb{R})$. La fibre type $\mathbb{R}^n$ est l'espace homogène $\mathcal{A}(n, \mathbb{R})/GL(n, \mathbb{R})$, où $\mathcal{A}(n, \mathbb{R})$ est le groupe affine. On en déduit que $H(M)$ est muni de connexions de Cartan ; à chaque forme de connexions ω (à valeurs dans $gl(n, \mathbb{N})$ est associée la forme de connexion de Cartan $\omega+\theta$ où θ est la forme sur $H(M)$ définie par $\theta(Y_y) = y^{-1}T\pi(Y_y)$, où π est la projection $H(M) \to M$ et $y \in H(M)$. On vérifie que θ est à valeurs dans $\mathbb{R}^n$ et que l'algèbre de Lie de $\mathcal{A}(n, \mathbb{R})$ est $gl(n, \mathbb{R}) \oplus \mathbb{R}^n$.

Pour une G-structure définie par un G-sous-fibré principal $H_G(M)$ de $H(M)$, la restriction θ_G de θ à H_G est la forme fondamentale de la G-structure. On peut alors formuler le *problème d'équivalence locale* d'Élie Cartan concernant les G-structures sous la forme suivante : étant données sur des variétés M et M', de même dimension, des G-structures H_G et H'_G, de formes fondamentales θ_G et θ'_G, ces structures seront dites localement équivalentes si pour tout $x \in M$, il existe un difféomorphisme d'un voisinage ouvert V de x sur un ouvert de M' tel que l'image réciproque $f^*_{\theta'_G}$ soit la restriction à V de la forme θ_G. Ce problème a été étudié par l'auteur de cet exposé, dans sa thèse [36]

3.3 Pseudogoupes, jets, groupoïdes de Lie

1. La considération des "groupes infinis" d'É. Cartan d'une part, la considération des atlas définissant une structure de variété différentiable d'autre part a conduit C. Ehresmann à introduire la notion de *pseudogroupe de transformations*, généralisant les travaux de Veblen et J. H. C. Whitehead. En résumé, un pseudogroupe de transformations Γ sur un espace topologique A est un ensemble d'homéomorphismes dont la source et le but sont des ouverts de A satisfaisant les conditions suivantes :
   a) Si f appartient à Γ alors f^{-1} appartient à Γ
   b) Si f et f' éléments de Γ sont composables, alors ff' appartient à Γ.

c) l'application identique de A appartient à Γ
d) Si $V = \cup_{i\in I} V_i$, où les V_i sont des ouverts alors f, de source V_i appartient à Γ si f est une bijection et si la restriction f_i de f à chaque V_i appartient à Γ ; la condition est nécessaire et suffisante.

Soit dans $\mathbb{R}^n$ les pseudogroupes Γ^0 (resp. Γ^r, Γ^∞) de tous les homéomorphismes locaux (resp. de tous les difféomorphismes locaux de classe C^r, de classe C^∞). Un atlas de $\mathbb{R}^n$ vers un ensemble M compatible avec l'un de ces pseudogroupes (c'est-à-dire dont les changements de cartes locales appartiennent au pseudogroupe) définit sur M une structure de variété topologique (resp. différentiable de classe C^r, indéfiniment différentiable).

Depuis son exposé au Colloque de Topologie (Paris, 1947), C. Ehresmann établit les fondements de la Géométrie différentielle à partir des pseudogroupes.

Par exemple une *structure feuilletée* sur une variété M peut être définie par un atlas de $\mathbb{R}^p \times \mathbb{R}^q$ vers M compatible avec le pseudogroupe Γ des applications de la forme :

$$(x, y) \mapsto (\varphi(x), \phi(x, y))$$

où φ est un diffémorphisme local dans $\mathbb{R}^p$ et ϕ_x tel que $\phi_x(y) = \phi(x, y)$ est un diffémorphisme local dans $\mathbb{R}^q$.

Cette structure feuilletée sur M peut être définie par la donnée d'un champ complètement intégrable d'éléments de contact de dimension q.

De même un espace fibré E, de base B, fibre type F sur laquelle un groupe de Lie G opère effectivement, est défini par un atlas de $B \times F$ vers E compatible avec le pseudogroupe des applications :

$$(x, y) \mapsto (x, s_x y)$$

définies sur $V \times F$, telles que l'application $x \to s_x$ soit une application différentiable de l'ouvert V dans G.

Depuis 1952, C. Ehresmann a plus généralement développé la théorie des *structures locales* à partir des pseudogroupes. En utilisant la notion de *jet local*, il démontre de manière élégante un résultat obtenu de façon plus compliquée en 1936 : un espace localement homogène de Lie compact et simplement connexe est isomorphe à un espace homogène.

2. Le souci d'établir un calcul différentiel intrinséque sur les variétés déjà apparent dans son article sur les connexions, ainsi que l'éclaircissement de la notion de prolongement dans le travaux d'Élie Cartan ont conduit C. Ehresmann à introduire un autre outil fondamental en Géométrie différentielle, à savoir la notion de *jet infinitésimal*.

C. Ehresmann a exposé sa théorie des jets et des prolongements dans une série de Notes aux Comptes-rendus publiées en 1951 et 1952, puis dans une conférence au Colloque international de Géométrie différentielle de Strasbourg en 1953. Dans ces articles, C. Ehresmann a aussi développé sa théorie des pseudogroupes de Lie.

Un jet d'ordre r, de source $x \in M$, de but $y \in W$ (où M et W sont des variétés différentiables de classe C^s, avec $s \geq r$) est la classe d'équivalence d'applications de classe r d'un voisinage ouvert de x, à valeurs dans W, appliquant x sur y, et dont les dérivées partielles à l'aide de coordonnées locales admissibles prennent les mêmes

valeurs en x pour tous ordres $\leq r$. À la composition d'applications, correspond une composition de leurs jets. On peut définir dans le même cadre un *jet local*, classe d'équivalence d'applications coïncidant dans un voisinage de x.

La notion de jet donne en particulier celle de vitesse d'ordre r (à l'ordre 1 on retrouve les vecteurs tangents) et celle de repère d'ordre r. L'ensemble des repères d'ordre r est un fibré principal $H^r(M)$, de groupe structural L^r_n, extension du groupe $L^1_n = GL(n, \mathbb{R})$. Un *prolongement d'ordre r* de M est un espace fibré associé à $H^r(M)$.

L'ensemble $J^r(M, W)$ des r-jets de M dans W est un espace fibré. La notion de *système différentiel d'ordre r* se dégage naturellement : un tel système (qui est un système d'équations aux dérivées partielles global) est défini par une sous-variété de $J^r(M, W)$.

La notion de jet infinitésimal, à la fois si simple et si importante a permis de mieux poser les problèmes fondamentaux de la Géométrie différentielle locale ou globale. Elle a notamment permis à R. Thom de développer sa théorie des singularités à la suite des travaux de H. Whitney.

3. L'introduction des *groupoïdes* en géométrie différentielle a été initiée par C. Ehresmann dans son article sur les connexions d'abord en considérant le groupoïde des chemins fermés, ensuite en associant à tout fibré pincipal, de groupe structural G, l'ensemble désigné par $\Phi = PP^{-1}$ des isomorphismes de fibré sur fibre ; Φ est appelé maintenant le *groupoïde de jauge* de P. Il joue un grand rôle en Physique.

   Rappelons qu'un groupoïde Φ est un ensemble muni d'une loi de composition partiellement définie telle que :

   $1^0)$ pour tout $x \in \Phi$, il existe dans Φ une unité à droite $\alpha(x)$ et une unité à gauche $\beta(x)$ telles que $x \cdot \alpha(x) = x$ et $\beta(x) \cdot x = x$.

   $2^0)$ xy est défini si et seulement si $\alpha(x) = \beta(y)$. Si xy et yz sont définis, il en est de même de $(xy)z$ et $x(yz)$; alors $(xy)z = x(yz)$.

   $3^0)$ tout $x \in \Phi$ admet un inverse x^{-1} tel que

   $$x^{-1}x = \alpha(x), \quad xx^{-1} = \beta(x).$$

   On obtient ainsi deux projection $\alpha : \Phi \to \Phi_0$ et $\beta : \Phi \to \Phi_0$ où Φ_0 est le sous ensemble des unités (ou *base*). Le groupoïde est transitif si $\alpha \times \beta : \Phi \to \Phi_0 \times \Phi_0$ est surjectif.

   Un *groupoïde différentiable* est une variété *différentiable* Φ munie d'une structure de groupoïde telle que :

   $1^0)$ Φ_0 est une sous -variété de Φ.

   $2^0)$ Les applications α et β sont des submersions.

   $3^0)$ l'application $\mu : \Phi_2 \to \Phi$ (où $\Phi_2 \subset \Phi \times \Phi$ est l'ensemble des couples composables) est différentiable.

   $4^0)$ l'application $x \to x^{-1}$ de Φ sur Φ est différentiable. C. Ehresmann a réservé le nom de *groupoïde de Lie* aux groupoïdes différentiables transitifs (ils sont alors localement des produits) alors qu'actuellement on désigne par groupoïde de Lie tout groupoïde différentiable.

   Le groupoïde de jauge est un groupoïde transitif dont l'ensemble des unités s'identifie à la base M du fibré principal. Le groupoïde $\Phi = PP^{-1}$ est encore le quotient de $P \times P$ par l'action diagonale de G. En prolongeant l'application $P \times P \to \Phi$ en

une application $TP \times TP \to T\Phi$ que l'on restreint à la diagonale de $P \times P$ et aux couples dont le deuxième vecteur est nul, on obtient le fibré vectoriel, de base M, *des déplacements infinitésimaux des fibres* ; un élément d'un tel fibré est encore un vecteur α-vertical dont l'origine est une unité. Si P est l'espace homogène $\hat{G}/G$, on retrouve les déplacements infinitésimaux du repère mobile introduits par É. Cartan (voir dans la première partie, le paragraphe III 2). Plus généralement on peut associer à tout groupoïde de Lie le fibré des vecteurs α-verticaux le long de la variété des unités. Ce fibré est un *algébroïde de Lie* au sens de J. Pradines [48], [56]. Celui-ci a défini un algébroïde de Lie comme un fibré vectoriel $\mathcal{A} \to M$ dont l'espace des sections est muni d'une structure d'algèbre de Lie, admettant un morphisme $P : \mathcal{A} \to TM$ d'espaces vectoriels induisant un morphisme d'algèbre de Lie des sections de $\mathcal{A}$ dans les sections de TM ; ce morphisme est appelé *l'ancrage* (en anglais "anchor map"). Enfin le crochet des sections de $\mathcal{A}$ satisfait une relation de Leibniz. L'algébroïde de Lie d'un groupoïde de Lie joue le même rôle que l'algèbre de Lie d'un groupe de Lie ; mais il existe des algébroïdes de Lie qui ne sont pas associés à des groupoïdes de Lie.

À la suite des travaux de C. Ehresmann, la théorie des groupoïdes de Lie a donné lieu à de nombreuses études (J. Pradines, P. Libermann, K. Mac Kenzie). Surtout cette théorie a pris une grande extension ces douze dernières années à la suite de l'introduction par A. Weinstein [68] et indépendamment par M. V. Karasev [25] des groupoïdes symplectiques. Les travaux de A. Weinstein, C. Albert et P. Dazord ainsi que de leurs collaborateurs ont montré l'importance de ces groupoïdes en liaison avec les structures de Poisson et les problèmes de quantification. Voir la 3$^{\text{ème}}$ partie de cet article.

4. Avec les notions de groupoïde, de pseudogroupe de transformations, de jet, on a les outils nécessaires pour introduire les *pseudogroupes de Lie* ; ceux-ci ont été définis par C. Ehresmann pour donner une version globale des "groupes infinis" d'É. Cartan.

   Dans ce paragraphe, pour simplifier, nous supposerons les données différentiables de classe $\mathbb{C}^\infty$.

   Pour tout pseudogroupe Γ de difféomorphismes locaux sur une variété différentiable M, l'ensemble $J^q(\Gamma)$ (resp. $J^\wedge(\Gamma)$) des q-jets (resp. des jets locaux) des applications appartenant à Γ est un groupoïde différentiable. En particulier l'ensemble $\pi^q(M)$ des q-jets de tous les difféomorphismes locaux est le groupoïde de jauge du fibré $H^r(M)$des repères d'ordre q.

   Un sous-groupoïde Φ^q de $\pi^q(M)$ constitue un système différentiel d'ordre q ; une solution de Φ^q est un difféomorphisme local dont tous les q-jets appartiennent à Φ^q. Le groupoïde Φ^q est dit complètement intégrable si pour tout $y \in \Phi^q$, de source x, il existe une solution f telle que $j^q_x f = y$.

   Un *pseudougroupe de Lie* d'ordre q sur une variété M est un pseudogroupe Γ de difféomorphismes locaux tel que :

   a) pour $s = 0, \ldots, q$, $J^s(\Gamma)$ est une sous-variété de π^s(M).

   b) Γ est *complet* c'est-à-dire Γ est l'ensemble des solutions de $J^q(\Gamma)$.

   On distingue les pseudogroupes de *type fini*, de *degré* r, s'il existe $r \geq q$ tel que la projection de $J^r(\Gamma)$ sur $J^{r-1}(\Gamma)$ soit localement bijective ; $J^r(\Gamma)$ est alors défini localement par un système d'équations aux dérivées partielle de Mayer-Lie (voir les "groupes infinis" d'É. Cartan). On démontre qu'alors $J^s(\Gamma)$ et par suite $J^\wedge(\Gamma)$ sont

isomorphes à $J^r(\Gamma)$ pour tout $s >$r. Les autres pseudogroupes de Lie sont dits de *type infini* (ils correspondent aux "groupes infinis" d'É. Cartan).
Par exemple le pseudogroupe des automorphismes locaux d'une structure riemannienne est de type fini. Celui des automorphismes locaux d'une structure symplectique est de type infini.
Le développement des idées de C. Ehresmann permet de démontrer des théorèmes d'É. Cartan sur les "groupes infinis" : prolongements (en particulier prolongement en un pseudogroupe d'ordre 1), équations de "structure", sous-pseudogroupes de Lie. Voir P. Libermann. [36][37].
C. Ehresmann a démontré (Comptes-rendus, 1958, 360-362) les deux théorèmes suivants relatifs aux pseudogroupes de Lie de type fini.

a) Un pseudogoupe de Lie Γ, de type fini, sur une variété compacte simplement connexe se déduit par localisation d'un groupe de Lie G (c'est-à-dire les applications de Γ sont des restrictions d'applications définies par l'action de G).
b) le plus grand groupe de transformations globales contenu dans un pseudogroupe de Lie de type fini est un groupe de transformations de Lie.

Ceci généralise les résultats connus concernant les structures riemanniennes ou le parallélisme.
Les démonstrations nécessitent l'utilisation de la notion de *pseudogroupe infinitésimal* (notion introduite indépendamment par C. Ehresmann et par moi-même [38]). Un pseudogroupe infinitésimal E associé à un pseudogroupe de Lie est un ensemble de champs de vecteurs engendrant des groupes locaux à un paramètre appartenant à Γ et tels que pour deux champs de vecteurs X et X', de sources U et U', les champs $aX + bX'$ (où a et b sont des constantes) et $[X, X']$ de source $U \subset U$' appartiennent à E. Cette notion a été reprise ultérieurement par Singer et Sternberg, sous le nom de L.A.S.
Depuis 1960, l'étude des pseudogroupes de Lie a pris un essor considérable avec les travaux de Matsushima et Kuranishi, puis de Spencer ; celui-ci a introduit des méthodes cohomologiques liées à la théorie des faisceaux. À la suite des travaux de Spencer, on peut citer les travaux de Goldschmidt, Guillemin, Guillemin-Sternberg, Kumpera, Qué, Malgrange. Enfin P. Molino et C. Albert ont publié un ouvrage sur ce sujet [3]. Ces théories algébriques permettent de clarifier la notion d'involution.

3.4 Remarques

C. Ehresmann partageait très généreusement ses idées (publiées ou non) avec d'autres mathématiciens (non nécessairement ses élèves) à l'occasion de conversations privées ou dans ses cours. Ces idées sont simples et naturelles ; elles deviennent si indispensables que leur auteur a été parfois oublié ou imité !

C. Ehresmann fut un "grand" directeur de thèses. Ses premiers élèves furent J. Feldbau et G. Reeb venus de Strasbourg et réfugiés comme lui à l'Université de Clermont–Ferrand pendant la guerre. Le premier travailla sur les espaces fibrés et l'homotopie ; sa thèse était presque achevée lorsqu'il fut arrêté et déporté en 1943 ; il mourut en 1945, à peine libéré des camps de concentration. Le deuxième fit acte de pionnier dans la théorie des feuilletages ; une note aux Comptes-rendus écrite en collaboration avec C. Ehresmann, publiée en juin

1944, contenait déjà de nombreux résultats.

C. Ehresmann eut de nombreux élèves, certains venus de l'étranger comme Wu Wen Tsun (chinois), Kumpera et de Barros (brésiliens). Les mathématiciens japonais Kobayashi et Nomizu commencèrent leurs recherches à Strasbourg avant d'émigrer aux États-Unis. Enfin A. Haefliger (suisse) travailla avec C. Ehresmann à Strasbourg puis à Paris. Le dernier élève de C. Ehresmann en Géométrie différentielle fut J. Pradines. Pendant la préparation de ma thèse à Strasbourg (de 1947 à 1953), j'ai eu la chance d'assister à l'élaboration et à la maturation des idées dégagées par C. Ehresmann en Géométrie différentielle.

Dans ses cours de 3^e cycle, C. Ehresmann ne faisait peut-être pas preuve de qualités pédagogiques, au sens classique du terme. En effet, ses cours n'étaient pas toujours très "structurés", des idées nouvelles venant interférer avec ce qui avait été préparé ; pour les auditeurs de tels exposés étaient très enrichissants.

R. Thom a été très influencé par les cours donnés par Ehresmann durant la période 1947-1951 (fondements de la Topologie algébrique et de la Géométrie différentielle).

De même un cours professé à Paris en 1957-1958 a été une source d'inspiration pour de nombreux mathématiciens qui étaient directement ses élèves (A. Haefliger, J. Pradines, P. Ver Ecke, C. de Barros, A. Kumpera, F. Benzecri) ou des élèves de A. Lichnerowicz (P. Molino, D. Lehmann, J. Lejeune-Lehmann, R. Bkouche, Que) ou encore des élèves de H. Cartan (J. P. Benzecri, Shih Weishu). En particulier A. Haefliger, dans sa thèse [23] a approfondi les idées de C. Ehresmann concernant les structures feuilletées et les groupoïdes.

C. Ehresmann organisa d'abord à Strasbourg de 1945 à 1954, puis à Paris à partir de 1955, un Séminaire où les jeunes mathématiciens exposaient leurs travaux et rencontraient des mathématiciens étrangers. C'est ainsi que vinrent à Strasbourg les mathématiciens G. de Rham (suisse), N. Kuiper (néerlandais), Hirzebruch, Dold, Stein, Remmert (allemands). Après les épreuves de la guerre, ces rencontres contribuèrent au rapprochement des mathématiciens de tous pays.

4 Troisième partie : Sur certains travaux de A. Lichnerowicz (1915-1998)

A. Lichnerowicz fut à la fois intéressé par la Relativité (sa thèse sous la direction de G. Darmois fut consacrée à ce sujet) la Mécanique, la Géométrie différentielle (connexions et groupe d'holonomie, Géométrie riemannienne, Géométrie complexe avec étude des variétés kählériennes). Son œuvre est marquée par l'influence d' É. Cartan. Deux ouvrages de géométrie différentielle ont été publiés par A. Lichnerowicz. Ce sont *Théorie globale des connexions et les groupes d'holonomie* [43] (dont la première partie a été inspirée par l'article de C. Ehresmann sur les connexions) et *Géométrie des groupes de transformations* [44].

La *Gazette des Mathématiciens* de la S. M. F. a publié dans son n ° 82 (octobre 1999) des analyses de l'œuvre de Lichnerowicz. Ce sont les articles de Marcel Berger (Lichnerowicz et la Géométrie différentielle), Yvonne Choquet-Bruhat (Lichnerowicz et la Relativité générale), Charles-Michel Marle (l'œuvre de A. Lichnerowicz en géométrie symplectique), ainsi qu'un article d'André Revuz sur la réforme des Mathématiques. Ces articles ont été traduits en Anglais et publiés dans les "Notices" de l'A. M. S. (décembre 1999) ; cette dernière publication

contient en outre un article de Jean-Pierre Bourguignon sur l'ensemble de l'œuvre.

En traitant des variétés de Poisson et de Jacobi nous utiliserons notamment l'article de C. Marle [52].

4.1 Sur les structures de Poisson

L'importance des crochets de Poisson en Mécanique classique a été reconnue par Lagrange , Poisson, Jacobi, Hamilton, Sophus Lie. Dans son livre sur les Invariants Intégraux, É. Cartan définit le crochet de Poisson et le crochet de Jacobi au moyen de l'algèbre extérieure. Depuis une cinquantaine d'années, à la suite notamment des travaux de A. Lichnerowicz, G. Reeb, J. M. Souriau, l'étude des problèmes de Mécanique au moyen de structures symplectiques et de contact a pris un nouvel essor mais le cadre de ces structures s'est avéré trop étroit. Certains auteurs tel J. M. Souriau ont utilisé les structures présymplectiques (définies par une 2–forme fermée de rang quelconque) mais c'est A. Lichnerowicz [46] qui le premier a vu que dans beaucoup de problèmes le *point de vue contravariant* convenait mieux que le point de vue covariant. Déjà É. Cartan avait associé au crochet de Poisson une "2-forme quadratique". C'est A. Lichnerowicz qui a clarifié la situation en montrant que le crochet était en réalité défini par un champ de bivecteurs.

Rappelons que sur une variété symplectique (M, Ω), la dualité transforme Ω en un champ de bivecteurs Λ tel que le crochet de Poisson $\{f, g\}$ de deux fonctions différentiables soit égal à $\Lambda(df, dg)$.

Plus généralement A. Lichnerowicz [46] a défini les *structures de Poisson* par la donnée d'un champ de bivecteurs Λ, de rang quelconque, de façon que le crochet $\{f, g\} = \Lambda(df, dg)$ satisfasse l'identité de Jacobi. Il a montré que cette condition est équivalente à $[\Lambda, \Lambda] = 0$, où $[\,]$ désigne le crochet de Schouten. Le champ Λ sera plus brièvement désigné par *bivecteur de Poisson.*

Auparavant A. Lichnerowicz [45] avait été amené à considérer les *variétés canoniques*, cadre approprié pour étudier du point de vue global les transformations canoniques de la Mécanique. Une variété M, de dimension $2n + 1$, est munie d'une structure canonique par la donnée d'un bivecteur de Poisson, de rang $2n$, et d'une application $t : M \to \mathbb{R}$ telle que le crochet de Schouten $[\Lambda, t]$ s'annule.

A. Lichnerowicz a montré que si une variété M, de dimension m, admet un tenseur de Poisson, de rang constant $2p < m$, le noyau de Λ est un système différentiel complètement intégrable définissant un feuilletage dont les feuilles sont des variétés symplectiques de dimension $2p$. Si $m = 2p$, la variété M est symplectique. L'auteur a démontré qu'il existait au voisinage de chaque point une variété de Poisson, des coordonnées locales "canoniques" $x^1, \ldots, x^p, y^1, \ldots, y^p, z^1, \ldots, z^{m-2p}$ telles que : $\{x^i, x^j\} = \{y^i, y^j\} = \{x^i, z^k\} = \{y^i, z^k\} = \{z^k, z^l\} = 0$, $\{y^i, x^i\} = \delta_{ij}$ pour tous *i, j, k, l* satisfaisant $1 \leq i, j \leq p$, $1 \leq k, l \leq m \leq m-2p$.

Soit $\Lambda^\sharp : T^*M \to TM$ le morphisme de fibrés vectoriels tel que pour tout $x \in M$ et tout couple (α,β) d'éléments de $T^*_x M$ on ait : $< \Lambda^\sharp\alpha, \beta >= \Lambda(\alpha, \beta)$. À toute fonction f sur M, on associe le *champ de vecteurs hamiltonien* $X_f = \Lambda^\sharp df$. Le sous-espace vectoriel $\Lambda^\sharp(T_x M)$ appelé *champ caractéristique* de Λ est engendré par les champs hamiltoniens ; c'est le fibré tangent au feuilletage symplectique.

A. Lichnerowicz a introduit dans [46] la notion de cohomologie associée à une structure de Poisson, que l'on pourrait appeler *cohomologie de Lichnerowicz-Poisson.* Il a remarqué

que l'opérateur ∂_Λ qui associe à tout p-vecteur P le $(p+1)$–vecteur $\partial_\Lambda P = [\Lambda, P]$ (ce crochet étant le crochet de Schouten) est de carré nul. La cohomologie ainsi définie est en général compliquée. Elle est activement étudiée par de nombreux chercheurs. Voir l'ouvrage de Vaisman [66]. A. Lichnerowicz a démontré que pour toute p-forme différentielle sur M, on a $\Lambda^\sharp(d\eta) = \partial_\Lambda(\Lambda^\sharp\eta)$ déterminant ainsi un homomorphisme de la cohomologie de Rham dans la cohomologie de Lichnerowicz–Poisson ; dans le cas d'une variété symplectique, l'homomorphisme est un isomorphisme.

Les structures de Poisson considérées ci-dessus sont *régulières* car le bivecteur Λ est de rang constant.

L'importance des variétés de Poisson (non nécessairement régulières) a été rapidement reconnue, notamment par A. Weinstein [67] qui en a étudié les propriétés locales. En plus de Weinstein et de ses collaborateurs de nombreux mathématiciens se sont intéressé à ce sujet : P. Dazord et ses élèves, C. Albert, G. Hector, C. Marle, Y. Kosmann–Schwarzbach ; enfin V. Drinfeld a créé la théorie des groupes de Lie–Poisson.

Quand le rang de Λ n'est pas constant, le champ caractéristique n'est pas un sous-fibré vectoriel de TM mais il définit une distribution complètement intégrable au sens de Sussmann, engendrée par les champs hamiltoniens dont les feuilles sont symplectiques. La complète intégrabilité peut être démontrée au moyen d'un théorème de Sussmann. Dans [42], les auteurs en ont donné une démonstration directe en utilisant des coordonnées locales adaptées.

Le bivecteur de Poisson étant un tenseur contravariant, on peut définir les morphismes de Poisson $\varphi : (M_1, \Lambda_1) \to (M_2, \Lambda_2)$ tels que $\varphi^*\{f_2, g_2\}_2 = \{\varphi^* f_2, \varphi^* g_2\}_1$

Soit (M_1, Λ_1) une variété de Poisson et $\varphi : M_1 \to M_2$ une submersion surjective vérifiant la condition (a) : pour tout couple (f_1, g_1) de fonctions sur M_1, constantes sur chacune des fibres $\varphi^{-1}(x_2)$, le crochet $\{f_1, g_1\}_1$ est aussi constant sur chaque fibre. Alors il existe sur M_2 une structure une structure de Poisson unique telle que φ soit un morphisme de Poisson. Ceci s'applique notamment aux groupoïdes symplectiques Φ introduits par A. Weinstein ; les α-fibres et les β-fibres étant symplectiquement orthogonales, il résulte de la théorie des feuilletages symplectiquement réguliers [39] que l'application $\alpha : \Phi \to \Phi_0$ vérifie la condition (a) et la base Φ_0 est munie d'une structure de Poisson. On a la structure de Poisson opposée sur Φ_0 en considérant les β-fibres.

On démontre que le dual $\mathcal{A}^*$ d'un algéboïde de Lie $\mathcal{A}$ est muni d'une structure de Poisson homogène et réciproquement. D'autre part on démontre que fibré cotangent T^*M d'une variété de Poisson (M, Λ) est muni d'une structure d'algébroïde de Lie dont l'ancrage est l'application $\Lambda^\sharp : T^*M \to TM$. En particulier l'algébroïde de Lie associé à un groupoïde symplectique est isomorphe à $T^*\Phi_0$.

Ainsi les deux problèmes suivants sont liés

1^0) intégration d'un algébroïde de Lie (recherche d'un groupoïde de Lie associé).

2^0) réalisation d'une structure de Poisson (M, Λ) c'est–à-dire la recherche d'un groupoïde symplectique dont M serait la base.

On n'abordera pas la question des déformations de l'algèbre des fonctions sur une variété de Poisson pour laquelle une importante contribution est due à A. Lichnerowicz et ses collaborateurs. Ceux-ci ont développé une méthode de *quantification* des systèmes hamiltoniens classiques, distincte de la méthode basée sur la quantification géométrique de B. Kostant et J. M. Souriau [33], [60].

L'ouvrage de I. Vaisman [66] sur les variétés de Poisson est une étude très complète de ces sujets ; il contient une abondante bibliographie.

4.2 Sur les stuctures de Jacobi

A. Lichnerowicz a étudié du point de vue contravariant les structures Pfaffiennes (cf. [38]) définies sur une variété de dimension $2n+1$ par une forme de Pfaff ω telle que $\omega \wedge (d\omega)^n \neq 0$ en tout point. Auparavant les automorphismes infinitésimaux conformes de ω et le crochet de Jacobi ont été étudiés dans [38] et repris dans [42].

A. Lichnerowicz a introduit la notion de *variété de Jacobi ;* une telle structure est définie sur une variété Mpar le couple (Λ, E) d'un champ de bivecteurs Λ et d'un champ de vecteurs E satisfaisant

$$[E, \Lambda] = 0, \quad [\Lambda, \Lambda] = 2E \wedge \Lambda$$

où [,] désigne le crochet de Schouten. Alors le crochet de Jacobi défini par

$$\{f, g\} = \Lambda(df, dg) + < fdg - gdf, E >$$

satisfait la relation de Jacobi.

Les structures Pfaffiennes correspondent au cas de dim $M = 2n + 1$ et $E \wedge (\Lambda)^n \neq 0$. Le champ E est lors le champ de Reeb tel que $i(E)d\omega = 0$, $i(E)\omega = 1$.

Pour généraliser les structures de contact (définies sur une variété M de dimension $2n+1$ par un sous–fibré de rang 1 du fibré T^*M engendré localement par une forme de contact), A. Lichnerowicz a introduit les variétés *conformes de Jacobi* qui correspondent aux *fibrés de Jacobi.* Un tel fibré est un fibré vectoriel P, de base M, dont les fibres sont de dimension 1, tel que l'espace des sections différentiables est muni d'une loi de composition $(s_1, s_2) \to \{s_1, s_2\}$ qui en fait une algèbre de Lie, cette loi de composition étant locale (le support de $\{s_1, s_2\}$ est contenu dans l'intersection des supports de s_1 et s_2). C'est ce que A. Kirillov a appelé *algèbre de Lie locale.*

F. Guedira et A. Lichnerowicz [21] d'une part, C. Marle [50, 51, 52, 53] d'autre part ont montré l'existence sur P d'une structure de Poisson homogène relativement au champ de vecteurs opposé au champ de Liouville. Si l'on considère le dual P^* de P, le complémentaire de la section nulle est muni d'une structure de Poisson. Dans le cas d'une structure de contact, on retrouve la symplectifiée (au sens d'Arnold) de cette structure de contact (voir [42])

Pour une structure de Jacobi (M, Λ, E), on définit le *champ caractéristique* engendré par l'image du morphisme $\Lambda^\sharp$ et le champ E ; ce champ n'est pas en général de rang constant. A. Kirillov a montré que ce champ de directions est complètement intégrable au sens de Sussmann, définissant un feuilletage généralisé dont les feuilles sont symplectiques (si elles sont de dimension paire) ou des variétés pfaffiennes (si elles sont de dimension impaire).

P. Dazord, A. Lichnerowicz, C. Marle [17] ont fait une étude locale des variétés de Jacobi et des variétés de Poisson homogènes qui leur sont associées.

4.3 Sur l'enseignement d'A. Lichnerowicz

Dans ses cours (d'abord à l'Université de Strasbourg et à l'Université de Paris, puis au Collège de France) A. Lichnerowicz faisait preuve d'une grande précision dans les démonstrations

les plus compliquées. Ses auditeurs admiraient sa dextérité dans les calculs les plus complexes. Mais derrière cette technique se dégageaient des idées profondes et fécondes.

Son ouverture d'esprit dans un grand nombre de domaines se montra dans le choix varié des sujets proposés à ses élèves de recherches. A. Lichnerowicz fut un excellent directeur de thèse pour ses élèves qu'il encourageait. Parmi eux citons Y. Choquet-Bruhat, M. Berger, C. Marle, P. Molino, P. Dazord, R. Bkouche, D. Lehmann, J. Lejeune-Lehmann, Que, Yvette Kosmann-Schwarzbach, Y. et H. Kerbrat.

5 Conclusion

Bien qu'ayant des formes d'esprit différentes, C. Ehresmann et A. Lichnerowicz, influencés par la pensée d'É. Cartan, avaient en commun le goût de la Géométrie.

Un exemple de "l'unité de la Géométrie dans sa grande diversité" est celui des relations entre algébroïdes de Lie et structures de Poisson. Dans l'œuvre d'É. Cartan la méthode du repère mobile et ses déplacements infinitésimaux et le livre sur les "Invariants intégraux" paraissent éloignés. Or C. Ehresmann introduisant le groupoïde de jauge d'un fibré principal et les déplacements infinitésimaux des fibres est conduit à ce que J. Pradines a appelé les algébroïdes de Lie. D'autre part la lecture du livre d'É. Cartan sur les "Invariants intégraux" a été une des motivations d'A. Lichnerowicz pour introduire les structures de Poisson.

C. Ehresmann s'intéressait surtout aux structures formant les bases de la Géométrie différentielle ; A. Lichnerowicz s'intéressait aux applications de la Géométrie à la Mécanique, la Relativité et la Physique Mathématique.

Chacun d'eux a créé une école mathématique qui, grâce à leur qualité humaine, alliée à des qualités d'ouverture intellectuelle, est devenue une grande famille.

C. Ehresmann et A. Lichnerowicz ont organisé en commun un Colloque International de Géométrie différentielles (Strasbourg 1953). Ce colloque ainsi que ceux de Paris (1947), Bruxelles (1950), Venise–Pise–Bologne (1953) et le congrès international des Mathématiciens (Harvard 1950) ont marqué l'histoire des mathématiques. Ils ont permis à des mathématiciens séparés par la guerre ou le "rideau de fer" de se retrouver. A une époque de grande effervescence intellectuelle mais où n'existaient ni fax, ni Internet et commençait seulement à se développer l'usage de textes multigraphiés, ces réunions de mathématiciens se montrèrent d'une grande utilité. Chacun de ces Colloques apportait des théories nouvelles ou des résultats nouveaux.

Bibliographie

[1] En ce qui concerne É. Cartan (à l'exception de ses livres) et C. Ehresmann nous renvoyons aux œuvres complètes.

[2] C. Albert et P. Dazord, *Groupoïdes de Lie et Groupoïdes symplectiques*, Séminaire Sud-Rhodanien à Berkeley (1989). M. S. R. I Publ. 20, Springer-Verlag, Berlin-Heidelberg-New York (1991), 1-12. Publications Dept. Math. Lyon 1988, 53-105 ; 1990, 27-99.

[3] C. Albert et P. Molino, *Pseudogroupes de Lie transitifs*. Travaux en cours, Hermann, Paris, 1987.

[4] D. Bernard, *Sur la géométrie différentielle des G-structures*, Ann. Inst. Fourier 10 (1960) 151-270.

[5] D. Bernard, *Espaces homogènes et repère mobile*, 4^e coll. Int. Géom. Diff. Santiago de Compotela, 1978, 55-63.

[6] R. Bkouche,*Lire Cartan dans les années 50 et 60 : la théorie des structures infinitésimales*, Conférence à ce Colloque.

[7] É.Cartan, *Leçons sur les invariants intégraux*, Paris, Hermann, 1922.

[8] É.Cartan, *La théorie des groupes finis et continus et la Géométrie différentielle*, traitées par la méthode du repère mobile (rédigée par J. Leray), Paris, Gauthier-Villars, 1937, 2^e édition 1951.

[9] É.Cartan et A. Einstein,*Letters on absolute parallelism*, 1929-32 Princeton University Press et Académie Royale de Belgique, Princeton, 1979.

[10] É. Cartan, *Les systèmes différentiels extérieurs et leurs applications géométriques*, Paris, Hermann, 1945.

[11] É. Cartan, *Œuvres complètes ; Partie I*, Paris, Gaulthier-Villars, 1952,
Partie II, Paris, Gaulthier-Villars, 1953, 2^e édition CNRS 1984.

[12] S. S. Chern, *Pseudogroupes continus infinis*, Colloq. de Géom. Diff. Strasbourg 1953, publications CNRS 119-136.

[13] A. Coste, P. Dazord et A. Weinstein, *Groupoïdes symplectiques*, Publ. Dept. Math. Lyon 2/A (1987), 1-62.

[14] P. Dazord, *Stabilité et linéarisation dans les variétés de Poisson, Géométrie symplectique et mécanique* (J. P. Dufour éd.), Hermann, Paris, 1985, 59-75.

[15] P. Dazord, *Groupoïdes symplectiques et troisième théorème de Lie*, Géométrie symplectique et Mécanique, Proceedings, 1988 (C. Albert éd.), Lecture de Notes in Math., 14-16, Springer-Verlag, Berlin-Heidelberg-New York, 1990, 39-74.

[16] P. Dazord et G. Hector, *Intégration symplectique des variétés de Poisson totalement asphériques*, Séminaire Sud-Rhodanien à Berkeley (1989). M. S. R. I Publ. 20, Springer-Verlag, Berlin-Heidelberg-New York (1991), 37-72.

[17] P. Dazord, A. Lichnerowicz et C. Marle, *Structure locale des variétés de Jacobi*, J. Math. Pures et appl. 70 (1991), 101-152.

[18] P. Dazord et D. Sondaz, *Variétés de Poisson et Algébroïdes de Lie*, Publ. Math. Univ. Lyon 1, 1/B (1988), 1-68.

[19] C. Ehresmann, *Œuvres complètes et commentées, Parties I_1 et I_2, Topologie algébrique et Géométrie différentielle*. Supplément des cahiers de Top. Alg. Et Géom. Diff., Amiens, 1984.

[20] M. Flato, A. Lichnerowicz et D. Sternheimer, *Déformations 1-différentiables des algèbres de Lie attachées à une variété symplectique ou de contact*, Compositio Math. 31 (1975) 47-82.

[21] F. Guédira et A. Lichnerowicz, *Géométrie des algèbres de Lie de Kirillov*. J. Math. Pures et appl. 63 (1984), 407-484.

[22] V. Guillemin et S. Sternberg, *Deformation theory of pseudogroup structures*, Mem. A. M. S. 64 (1966).

[23] A. Haefliger, *Structures feuilletées et cohomologie à valeur dans un faisceau de groupoïdes*, thèse, 1958, Commentarii Math. Helvetici, vol. 32, fasc 4, 1958.

[24] A. Haefliger, *Naissance et développement de la théorie des feuilletages*, conférence à ce Colloque.

[25] M. V. Karasev, *Analogues of objets of the Lie group theory for non linear Poisson brackets*, Soviet Math. Izvestia 28 (1987), 497-527.

[26] M. V. Karasev et V. P. Maslov, *Non linear Poisson brackets*. Geometry and quantization, transl. of Math. Monographies, A. M. S, 1993.

[27] A. Kirillov , *Local Lie algebras*, Russian Math. Surveys 31 (1976), 55-75.

[28] S. Kobayashi et K. Nomizu, *Foundations of Diff. Geom. I, II*, Intersci. Publ, New York-London, 1963, 1969.

[29] Y. Kosmann-Schwarbach, *Poisson-Drinfeld groups, Topics in Soliton theory* (M. Ablowitz, B. Fuchssteiner and Krusal eds) World scientific, Singapore, 1987, 191-215.

[30] Y. Kosmann-Schwarbach, *Exact Gerstenhaber algebras and Lie bialgebroids*, Acta Appl. Math, 41 (115), 153-165.

[31] Y. Kosmann-Schwarbach, *The Lie bialgebroid of a Poisson-Nijenhuis manifold*, Letters in Math. Phys 38 (1996) 421-428.

[32] Y. Kosmann-Schwarbach et F. Magri, *Poisson-Nijenhuis structures*, Ann. Inst. Henri Poincaré Phys. Théor. 53 (1990) 35-81.

[33] B. Kostant , *Quantization and unitary representations, Lectures in modern analysis and applications III.* Lecture Notes in Math. 170, Springer-Verlag, Berlin-Heidelberg-New york, 1970, 87-207.

[34] A. Kumpera – D. C. Spencer, *Lie equations*, Ann. of Math. Studies 73, Princeton Univ. Press, Princeton, 1972.

[35] J. L. Koszul, *Crochet de Schouten-Nijenhuis et cohomologie, É. Cartan et les Mathématiques d'aujourd'hui*, Soc. Math. de France, Astérix, hors série, 1985, 257-271.

[36] P. Libermann, *Sur le problème d'équivalence de certaines structures infinitésimales régulières*, thèse, Strasbourg, 1953 ; Ann. Math. Pure Appl. 36 (1954), 27-120.

[37] P. Libermann, *Sur les pseudogroupes de Lie*, Coll. Top. Strasbourg 1954-55.

[38] P. Libermann, *Sur les automorphismes infinitésimaux des structures symplectiques et de contact*, Colloq. de Géom. Diff. Globale, Bruxelles, 1958.

[39] P. Libermann, *Pseudogroupes infinitésimaux*, Bull. Soc. Math. France, 87 (1959), 409-425.

[40] P. Libermann, *Problèmes d'équivalence et Géométrie symplectique*, IIIe rencontre de Géométrie de Schnepfenried, vol. 1, Astéristique, 107-108 (1983) 43-68.

[41] P. Libermann, Elie Cartan, *Travaux math. fasc. VIII*, 1996, Publ. Centre Univ. Luxembourg.

[42] P. Libermann et C. M. Marle, *Géométrie symplectique et bases théoriques de la Mécanique*, publ. Math. de l'Univ. Pris VII, vol. 21, t. I, II, III, IV, V, Paris (1986). Symplectic Geometry and Analytical Mechnics. D. Reidel Publ. Comp. Dordrecht-Boston, 1987.

[43] A. Lichnerowicz, *Théorie globale des connexions et des groupes d'holonomie*, Ed. Cremonese, Rome, 1955.

[44] A. Lichnerowicz, *Géométrie des groupes de tansformations*, Paris, Dunod 1958.

[45] A. Lichnerowicz, *Variétés symplectiques, Variétés Canoniques et systèmes dynamiques*, Topics in Diff. Géom., Acad. Press publ. 1976, New-York, San Francisco, London, 57-85.

[46] A. Lichnerowicz, *Les variétés de Poisson et leurs algèbres de Lie associés*, J. Diff. Geometry 12 (1977) 253-300.

[47] A. Lichnerowicz, *Les variétés de Jacobi et leurs algèbres de Lie associées*, J. Math. Pures et Appl. 57 (1978), 453-488.

[48] K. Mackensie, *Lie groupoids and Lie algebroids* in Diff. Geom. London Math. Lecture notes series 124, Cambridge Univ. Press, Cambridge, 1987.

[49] B. Malgrange, *Equations de Lie*, J. Diff. Geometry 6 (1972), 503-522.

[50] C. M. Marle, *Quelques propriétés des variétés de Jacobi*, Sem. Sud-Rhodanien (J. P. Dufour éd.) Travaux en Cours, Hermann, Paris (1985), 125-139.

[51] C. M. Marle, *On Jacobi Manifolds and Jacobi Bundles*, Sem. Sud-Rhodanien à Berkeley (1989). M. S. R. I Publ. 20, Springer-verlag, Berlin-Heidelberg-New York (1991) 227-246.

[52] C. M. Marle, *L'œuvre d'A. Lichnerowicz en géométrie Symplectique*, gazette des Mathématiciens. 81. S. M. F 1999. Notices A. M. S. 46, n ° 11.

[53] C. M. Marle, *Espace et temps physiques et description des systèmes mécaniques*, conférence à ce colloque.

[54] P. Molino, *Théorie des G-structures. Problème d'équivalence*. Lecture Notes in Math. Springer-Verlag Berlin-Heidelberg, New York 1978.

[55] J. Pradines, *théorie de Lie pour les groupoïdes différentiables* I. C. R. Acad. Sc. Paris 263 (1966) 907-910 ; II C. R. Acad. Sc. Paris 264 (1967) 254-248.

[56] J. Pradines, *Troisième théorème de Lie pour les groupoïdes différentiables* C. R. Acad. Sc. Paris 267 (1968).

[57] G. Reeb, *Sur certaines propriétés topologiques des variétés feuilletées*, Thèse Strasbourg 1949 ; Hermann, Paris, 1952.

[58] G. Reeb, *Sur certaines propriétés topologiques des trajectoires des systèmes dynamiques*, Mémoires Acad. Royale Belgique Sc 27, 9 (1952) 130-194

[59] G. Reeb, *Sur les espaces de Finsler et de Cartan*. Colloq C.N.R.S, Paris, 1953.

[60] J. M. Souriau, *Structure des systèmes dynamiques*. Dunod, Paris, 1969.

[61] D. C. Spencer, *Deformations of structures on manifolds defined by transitive continuous pseudogroups* Ann. of Math. 76 (1962), 306-345.

[62] H. J. Sussmann, *Orbits of families of vector fields and integrability of distributions*, Trans. Amer. Math. soc. 180 (1973), 171-188.

[63] W. M. Tulczyjew, *The Legendre transformation*, Ann. Inst. H. Poincaré 27 (1977) 101-114.

[64] W. M. Tulczyjew, *Poisson reductions*, Séminaire Sud-Rhodanien (J. P. Dufour éd.) Travaux en cours, Hermann, Paris (1985), 113-120.

[65] W. M. Tulczyjew, *Geometric formulations of physical theories*, Bibliopolis Napoli (1989).

[66] I. Vaisman, *Lectures on the Geometry of Poisson maanifolds*. Birkhäuser, Basek-Boston-Berlin 1994

[67] A. Weinstein, *Lectures on symplectic manifolds*, C.B.M.S Conf. Series in Math. 29. A. M. S Providence R. I. 1977.

[68] A. Weinstein, *The local structure of Poisson manifolds*. J. Diff. Geom. 18 (1983), 523-557.

[69] A. Weinstein, *Blowing up realizations of Heisenberg-Poisson manifolds* Bull. Sc. Math (1989), 381-406.

La géométrie et le principe d'idonéité : une relecture de Ferdinand Gonseth

Gerhard Heinzmann,
Département de Philosophie, LPHS—Archives H. Poincaré,
UMR 7117
Université de Nancy 2

Dans une première partie, je me propose de restituer le cadre systématique de mes propres réflexions qui justifient un retour à la pensée de Gonseth. Dans une deuxième partie, je discute l'idée centrale de Gonseth concernant la géométrie.

1 Qu'est-ce que la philosophie de la géométrie/des mathématiques ?

1. A première vue, on pourrait croire que depuis le début du XXe siècle, il n'y a plus à chercher la nature du rapport de la géométrie à la réalité, parce que ce rapport est brisé [[16], 168]. La "réalité" est remplacée par l'algèbre et l'analyse qui procurent l'une dans la géométrie algébrique et l'autre dans la géométrie différentielle la rigueur des aspects géométriques [[7], 39]. Mais il n'y a pas seulement l'algébraisation de la géométrie, il y a également le mouvement inverse : par exemple, en Théorie des probabilités, nous dit Maurice Fréchet, le calcul des intégrales multiples étaient long et pénibles jusqu'on "a eu l'idée de considérer les n variables comme les coordonnées d'un point dans l'espace à n dimensions et ces intégrales comme volumes et masses dans cet espace" [[3], 122]. Ce qui reste, d'une manière plus générale est d'une part un mouvement vers des espaces abstraits (algébriques, topologiques, etc.) et, d'autre part, un usage de termes et de contenus géométriques comme, par exemple, "points", "distance" ou "dimension" en tant que cadre langagier dans les domaines les plus divers (l'espace de Hilbert, de Banach, la distance de Hamming [nombre de couple (a_i, b_i) tels que a_i soit différent de b_i] importante pour la théorie de codage [[1], 5]).
   Le **mathématicien** suisse Ferdinand Gonseth semble alors peu au fait de l'évolution des idées géométriques lorsqu'il rédige en 1936 un ouvrage sur "Les Mathématiques et la réalité" et de 1945 à 1955 un ouvrage de 629 pages consacrés à "La Géométrie et

le problème de l'espace". En effet, dans ce dernier ouvrage, il développe péniblement l'édification axiomatique de la géométrie euclidienne et le passage aux géométries non-euclidiennes, toutes choses appartenant déjà à l'époque au bagage d'un étudiant des premières années [[14], 62].

Cependant, peut-être le **philosophe** Gonseth est-il plus actuel ? Car ne faut-il pas accorder au philosophe un certain délai qui sépare le savoir réfléchi du savoir vivant ? [[17], 80] Les géométries non euclidiennes et la physique relativiste ou quantique n'ont pas remises en cause la conception philosophique d'une connaissance qui serait adéquation à une réalité *apriori,* immuable, développement dont Gonseth nous livre la réflexion ? [[17], 68]

Certes, c'est à la fois juste et trivial. Mais en même temps il faut mentionner qu'un apriorisme n'était pas toujours défendu en philosophie. Ainsi, il ne serait pas à exclure que la leçon méthodique que nous apprend la physique est simplement une confirmation des idées philosophiques, soutenues bien antérieurement, par exemple, dans le Pragmatisme de Peirce.

2. Quel rôle doit-on attribuer à la philosophie des mathématiques ? — Commençons par dire ce qu'elle n'est pas : la philosophie des mathématiques n'est pas
   - un savoir pré-théorique (Dilthey) où l'on discute des opinions et des goûts
   - un savoir post-théorique (Wundt), qui s'épuise dans une épistémologie, c'est-à-dire dans une description de ce que le mathématicien fait (*Nachschwätzerei*)
   - un cabinet de curiosités où l'on traite à la fois les questions laissées en suspens par le mathématicien et les difficultés de principe inhérentes aux mathématiques, mais mieux traitées par le mathématicien.

   Au contraire, la philosophie des mathématiques devrait étudier les présuppositions conceptuelles et méthodologiques qui sont à la base des mathématiques "actuelles" : en d'autres termes, les questions traitées ne doivent pas être sans intérêts pour la pratique du mathématicien. Le philosophe ferait bien de ne pas se limiter à la question de la vérité de $2+2 = 4$ sans qu'il soit nécessaire de commenter les derniers résultats techniques. L'exemple adéquat est l'exemple "simple" qui, méthodiquement, se laisse transposer aux résultats plus sophistiqués.

   L'objet d'études est bien la pratique mathématique mais non en tant que "gospel" mais en tant que point de départ d'une critique scientifique qui se distingue de la pratique mathématique bien plus dans la perspective que par rapport au standard scientifique adopté [11]. Or, la pratique scientifique est tributaire des normes influencées par le progrès scientifique lui-même. On appelle cette attitude modeste de la philosophie la Thèse "naturaliste".

3. Je suis maintenant en mesure de formuler ma thèse principale : Par rapport à ces perspectives, le **philosophe** Gonseth, quasiment inconnu, ce mathématicien loin des recherches effectuées à son époque et d'ailleurs frappé de cécité, fait une figure bien honorable. Honorable par rapport à un des philosophes les plus connus du siècle, adhérent à notre thèse naturaliste, et dont la compréhension technique des sciences ne fut guère mise en doute. Je parle de Willard Van Orman Quine[1].

[1] Bien que j'aie indiqué à plusieurs reprises le rôle précurseur de Gonseth par rapport au holisme de Quine [cf. e.g. [8], 122 ou [9]], je fais dans la suite une large utilisation de l'excellent article de Michael Esfeld [2] qui a traité ce sujet par rapport aux quatre principes méthodologiques de Gonseth.

4. Quine s'oppose avec sa thèse naturaliste à une position classique que l'on peut résumer de la manière suivante : une proposition mathématique est acceptée (vraie) non pas par rapport à une description du monde, mais par rapport à la signification de la proposition.
   Quine a mis en difficulté cette position classique dans son célèbre article "Deux dogmes de l'empirisme". Dès lors, la distinction entre description et signification, représentation et interprétation ou synthétique et analytique est suspecte. Il en suit une double conséquence :
   - la différence entre le rôle du langage et le rôle du monde dans la détermination de la vérité d'une phrase n'est pas si abrupte et ne joue pas un rôle si important que l'on croyait (bien qu'il subsiste le cas tranché de la logique).
   - Si les mathématiques perdent leur statut purement analytique (l'échec technique d'une réduction est clair depuis Russell), le progrès en mathématiques devient une notion problématique parce qu'il n'est plus en principe indépendant de l'expérience future. On ne peut plus simplement maintenir la vue cumulative du progrès mathématique [[10], 519].

   La thèse positive de l'holisme sémantique et épistémologique défendu par Quine (a) et d'autres (b) se résume de la manière suivante :
   a) Une phrase n'a pas de signification indépendamment d'une théorie ; une phrase isolée ne peut être confirmée ou réfutée mais seulement une théorie.
      Dans la suite, on parlera de la sous-détermination de l'expérience. Si on l'entend par confirmation une confirmation par l'expérience sensible, cette thèse holiste, appliquée aux mathématiques en général et à la géométrie en particulier, est sûrement trop forte. Pour cette raison, on peut associer l'holisme à une thèse d'indispensabilité :
   b) En général, la partie mathématique entière (et non pas seulement une phrase isolée) d'une théorie scientifique n'a pas le statut d'une théorie vraie, mais seulement d'une théorie indispensable.

   En d'autres termes : on devrait admettre une distinction entre les parties d'une théorie qui sont vraies et les parties qui ne sont qu'utiles. Parfois, les parties utiles sont indispensables dans le sens que nous ne possédons pas une autre bonne théorie qui fonctionne sans elles. Il s'ensuit que les parties indispensables des mathématiques dans une théorie scientifique bien confirmée ne servent plus à établir leur vérité [[15], 232].
5. En réunissant maintenant les thèses naturalistes et holistes étendues on arrive à trois difficultés majeures, bien connues dans la littérature :
   1° Si nous n'acceptons que cette partie des mathématiques qui sont nécessaires pour nos meilleures théories scientifiques, nous n'acceptons que peu de mathématiques [[15], 233].
   2° Comment l'expérience peut-elle confirmer une théorie ?
      a) On constate le dilemme suivant : Si l'expérience n'est fondée que sur un événement sensoriel comment un tel événement peut-il être la *raison* d'un éventuel changement conceptuel et si l'expérience est conceptuelle (mathématiques), elle devrait être elle-même incluse dans l'holisme.
      b) L'holisme est caractérisé par une tension entre d'une part la confirmation d'une connaissance partagée par toute une communauté et une confirmation concernant une personne singulière [[2], 5-7].

3 ° Une bonne raison de considérer une théorie comme falsifiée n'est donnée que sous la condition que l'on dispose d'une meilleure théorie expliquant les raisons pour lesquelles l'ancienne théorie était moins bonne. En d'autres termes, une théorie isolée n'est pas un bon candidat pour une falsification [[12], 204]. Cet argument conduit pour ainsi dire un dédoublement de l'argument holiste à un niveau supérieur.

2 La solution de Gonseth

1. Je me propose maintenant de montrer que les thèses développées par Gonseth dans son modèle géométrique de la connaissance évitent les trois difficultés inhérentes à l'holisme exposé au paragraphe 1.5.
La réflexion de Gonseth porte, exemplifiée au paradigme de la géométrie, sur l'utilisation des éléments dits "empiriques" — "intuitifs" et "formels" — "théoriques" dans les mathématiques. Selon Gonseth, les développements en philosophie et en mathématiques prouvent l'insuffisance de deux approches opposées qui veulent, l'une, attribuer à l'intuition ou à l'empirique un rôle fondamental, et l'autre, l'exclure de l'entreprise mathématique. En se limitant ainsi, d'une part, à la *constitution* des objets mathématiques et, d'autre part, à *l'analyse* d'un formalisme, on néglige dans les deux cas la relation entre la détermination des objets et la garantie de la validité des énoncés sur ces objets.
À partir de ce cadre, l'idée centrale de Gonseth se résume ainsi : la connaissance géométrique doit être conçue comme "construction schématique". Dans cette construction, la structure formelle de la géométrie et des géométries est obtenue grâce à un "processus d'abstraction" à plusieurs niveaux. Chacun d'entre eux est caractérisé par l'existence d'un équilibre entre les aspects pragmatiques, empiriques et théoriques.
Semblable à Peirce, Gonseth s'oriente - dans sa philosophie scientifique - au standard méthodique des sciences exactes, ici la géométrie. Comme Peirce, il cherche une solution entre l'empirisme et le rationalisme et refuse de considérer avec Kant les théorèmes mathématiques comme énoncés synthétiques - a priori. Gonseth cherche la possibilité de faire dépendre les moyens théoriques de la connaissance d'une "idonéité" pratique.
La méthode d'explication qui en résulte se résume en 4 maximes :
le principe de dualité exige un dialogue dans lequel des "horizons" de 'réalité' et de 'conceptualité' se spécifient l'un par rapport à l'autre.
le principe de révisibilité exige l'ouverture vis-à-vis du progrès scientifique,
le principe de technicité est une sorte de contrebalance du principe de révisibilité : il érige en instance légitime de la révisibilité un esprit formé par la pratique des techniques efficaces concernant le problème en question.
le principe de solidarité est l'expression d'un holisme méthodique à tous les niveaux de la pensée.
C'est selon ces 4 maximes que Gonseth conçoit la genèse logique de la géométrie. Plongées dans le dialogue des chercheurs, les étapes de cette genèse sont appelées 'synthèses dialectiques' entre la 'réalité', le 'sujet' et la 'théorie'. Gonseth poursuit ainsi deux buts épistémologiques en mathématiques :
i) Donner un cadre "idéal" par rapport auquel le monde extérieur peut être structuré

ii) Arriver à une compréhension et un développement systématique des mathématiques introduites [[10], 530]. Le lien entre ces deux buts est un cheminement dialectique dépendant de paramètres sociologiques. Il s'ensuit que le critère pour le progrès mathématique consiste dans une interpolation maximale par rapports aux deux buts et la donnée sociologique [[10], 531].

On constate d'une manière immédiate que le principe de dualité résout la difficulté 2b et le principe de solidarité résout la difficulté 3 du paragraphe 1.5. La difficulté 2a trouve pour sa part la solution dans le fait que, pour Gonseth, le retour à une base non-conceptuelle est une illusion. Il refuse donc une information à l'état pur tels que les objets et leurs liens sont déjà des constructions schématiques, ou, selon une autre terminologie gonsethienne, des horizons de réalité ouverts aux précisions. Reste la première difficulté. Nous y allons revenir au prochain paragraphe qui traitera de la construction de la genèse logique de la géométrie.

2. Au niveau intuitif de la construction géométrique, les aspects de l'action (pragmatique), de l'aperception (empirique) et du parler (théorique) se confondent et les connaissances géométriques se justifient par une réalisation directe. Le conflit, s'il existe à ce niveau, "n'est donc pas entre notre intuition et une théorie douteuse, mais entre un premier schéma dont nous avons tellement pris l'habitude que nous lui conférons une existence" en quelque sorte intuitive, matérielle (géométrie euclidienne) et un autre schéma, celui d'une géométrie non-euclidienne, dont la nouveauté nous surprend [[5], 110].

   L'approche conventionnaliste de Poincaré de la genèse des géométries non-euclidiennes illustre parfaitement le prochain palier vers les géométries non-euclidiennes. La géométrie est le résultat d'une réflexion sur certaines actions (aspect pragmatique) idéalisées, appelées déplacements. L'étude de la structure de ces déplacements (aspect théorique) nous suggère une autre idéalisation, à savoir de passer au concept mathématique de groupes continus. En effet, les déplacements se composent d'après les mêmes lois que les substitutions d'un certain groupe G de transformations. Les groupes isomorphes à G seront à leur tour classés à l'aide de leurs sous-groupes qui se prêtent à y rapporter notre expérience déjà mathématisée (aspect empirique). Ainsi on obtient une caractérisation des groupes qui correspond aux géométries à courbure constante. Parmi ces groupes, nous choisirons finalement celui qui permet "l'affirmation de l'existence d'un sous-groupe *invariant* dont tous les déplacements sont échangeables et qui est formé de toutes les translations" (Poincaré 1921, 34). En d'autres termes, nous choisirons le groupe qui correspond à la géométrie euclidienne, parce que ses sous-groupes sont mieux suggérés par l'expérience mathématisée.

   Nous voyons que dans cette reconstruction de Poincaré le principe de technicité limite le choix des conventions effectuées : les choix qui mènent à une révision de principe de la géométrie euclidienne se font dans le cadre d'une approche synthétique de la géométrie, fondée sur la théorie des groupes de transformation.

   Jusqu'ici, "nous avons bien fait du déplacement la notion géométrique fondamentale, mais rien n'y est resté [...] de la vitesse" [[5], 125] ; il manque pour atteindre le prochain palier dans la construction gonsethienne, le temps. Le plus simple de représenter les relations de temps avec l'espace euclidien, c'est-à-dire entre les géométries de deux observateurs en mouvement, relativement l'un à l'autre, est de construire un modèle

géométrique à quatre dimensions. Le résultat bien connu est le suivant :
"Le problème des mouvements relatifs de l'espace euclidien met en lumière un groupe de 10 paramètres de transformations de l'univers en lui-même". Ce groupe est comparable à un groupe de déplacements non-euclidiens d'un espace à quatre dimensions d'une espèce fort dégénérée [[5], 127sq.]. "Les objets de la Mécanique classique apparaissent donc simplement comme des objets définis, grâce à une métrique non-euclidienne à quatre dimensions et de caractère très spécial ; le passage d'un système inertiel à un autre est un simple déplacement non-euclidien de cet espace sur lui-même". [[5], 131]
À ce stade, le principe de solidarité qui nous oblige à prendre en considération le temps, nous fait réviser la position de commodité poincaréienne. Il me semble remarquable que l'accord avec l'expérience est toujours maintenue. Il n'est détruit que par une autre application du principe de solidarité : "lorsqu'on veut assimiler les déplacements relatifs d'un observateur "galiléen" et d'une onde lumineuse à des déplacements galiléens" [[5], 146]. Pour le rétablir, il faut passer de la transformation de Galilée à celle de Lorentz. Ainsi, les modèles relativistes permettent d'interpréter les systèmes inértiaux aussi bien du point de vue de l'optique que du point de vue mécanique. Du point de vue relativiste la mécanique et l'optique ont la même cinématique.
Cependant, est-il *logiquement* nécessaire que la mécanique et l'optique obéissent aux mêmes lois de relativité ? [[5], 147] Évidemment non, il faut faire une nouvelle synthèse :

> "il n'y a aucune contradiction logique à accepter une cinématique des systèmes matériels, et une autre cinématique des systèmes ondulatoires. L'essentiel est de résoudre parfaitement le problème de relativité, qui se pose maintenant sous une triple forme : pour les systèmes matériels entre eux, pour les systèmes ondulatoires entre eux, et pour les relations des premiers avec les seconds. Ceci fait — et la solution est possible puisqu'il suffit de traduire l'einsteinien en galiléen — la théorie sera constituée de façon cohérente [...]. La construction des univers relativistes a définitivement brisé le caractère de tangible réalité qu'on attribuait à la cinématique classique comme à l'espace euclidien avant la construction des espaces non-euclidiens." [[5], 148]

J'ai commencé cet article par un constat semblable et pourtant bien plus imprécis : "il n'y a plus à chercher la nature du rapport de la géométrie à la réalité, parce que ce rapport est brisé". Ce qui est brisé jusqu'à présent n'est que la *tangible* réalité euclidienne. L'enseignement que je propose de tirer de la philosophie de Gonseth revient à dire que le constat d'une coupure absolue est en fait le constat d'une attitude qui méconnaît le problème des fondements en faisant l'économie de l'histoire et de l'évolution de la pensée abstraite. Il n'y a pas seulement une évolution des contenus mathématiques mais en même temps une évolution des moyens de la connaissance permettant d'inventer des synthèses nouvelles par rapport auxquelles la coupure n'existe pas : il n'y a pas de l'abandon de l'aspect intuitif parce que le référentiel intuitif n'est pas le référentiel (la réalité) initial(e). C'est de cette manière qu'il faut interpréter l'extension de l'espace et la géométrisation des contenus. Ainsi, par exemple, Jeremy Gray constate que lors de la création de la géométrie algébrique il y avait bien une unité d'aspects spécifiquement géométriques et d'aspects appartenant à la théorie complexe

des fonctions [[7], 43]. Philippe Nabonnand montre, pour sa part, dans un travail à publier, que l'approche synthétique des géométries à l'aide de groupe de transformations (donc le départ pour Poincaré et Gonseth) rencontre finalement la conception analytique de la géométrie, celle, plus riche, qui s'appuie sur la métrique et l'infinitésimal. Ces dernières, pour Poincaré encore qu'un chapitre d'Analyse, se trouvent unifiées aux théories émanant du point de vue synthétique par l'introduction de la notion de connexion par Cartan et Weyl . Ce qui paraît temporairement une autonomie de l'horizon abstrait doit pousser d'autres au travail d'unification et d'application aux domaines mathématiques ou scientifiques bien établis. Il faut laisser s'appliquer les mathématiques aux mathématiques. Ainsi se trouve finalement résolue la difficulté 1 ° du paragraphe 1.5.

Bibliographie

[1] N. Christmann, *Was ist (war) Geometrie ?*, in : *Skripten zur Fachdidaktik* 5, Fachbereich Mathematik, Universität Kaiserslautern, 5-19, 1996.

[2] M. Esfeld, *Gonseth and Quine,* Dialectica 55, à paraître, 2001.

[3] M. Fréchet, *De l'espace à trois dimensions aux espaces abstraits*, in : F. Le Lionnais (éd.), *Les grands courants de la pensée mathématique*, Paris : Blanchard, 121-129, 1962.

[4] F. Gonseth, *La géométrie et le problème de l'espace*, vol. I-VI, Neuchâtel : Griffon, 1945/55.

[5] F. Gonseth, *Les fondements des mathematiques. De la géometrie d'Euclide à la relativité générale et à l'intuitionnisme* ([1]1926), Paris : Blanchard, 1974.

[6] F. Gonseth, G.-H. Müller, *Philosophie mathématique* dans *Philosophie, chronique actuelle*, Paris : Hermann (ASI 1105), 33-127, 1950.

[7] J. Gray, *Algebraic Geometry between Nœther and Nœther — A Forgotten Chapter in the History of Algebraic Geometry*, Revue d'histoire des mathématiques 3, 1-48, 1997 .

[8] G. Heinzmann, *Schematisierte Strukturen. Eine Untersuchung über den Idoneismus Ferdinand Gonseths auf dem Hintergrund eines konstruktivistischen Ansatzes*, Bern/Stuttgart : Haupt, 1982.

[9] G. Heinzmann, *La Géométrie comme paradigme de l'idonéisme*, dans M. Panza/J.-C. Pont, *Espace et horizon de réalité. Philosophie mathématique de Ferdinand Gonseth*, Paris/Milan/Barcelone/Bonn : Masson 1992, pp. 47-54.

[10] P. Kitcher, *Mathematical Progress*, Revue Internationale de Philosophie 167, 518-540, 1988.

[11] P. Maddy, *Naturalizing Mathematical Methodology*, dans Schirn, Matthias (éd.) *The Philosophy of Mathematics Today*, Oxford : Clarendon Press, 175-193, 1998.

[12] F. Müllhölzer, *Science without Reference ?*, dans U. Majer/H-J. Schmidt, *Reflections on Spacetime. Foundations, Philosophy, History*, Dordrecht/Boston/London : Kluwer, 203-222, 1995.

[13] H. Poincaré, *Des fondements de la géométrie* ([1]The Monist 1898/99, pp.1-43), ed. L. Rougier, Paris : Chiron (s.d.), 1921.

[14] J.-C. Pont, *La géométrie et le problème de l'espace chez Gonseth* dans M. Panza/J.-C. Pont, *Espace et horizon de réalité. Philosophie mathématique de Ferdinand Gonseth*, Paris/Milan/Barcelone/Bonn : Masson 1992, 57-82, 1992.

[15] M. D. Resnik, *Holistic Mathematics* dans Schirn, Matthias (éd.) *The Philosophy of Mathematics Today*, Oxford : Clarendon Press, 227-246, 1998.

[16] H. Sinaceur, *De la Géométrie formelle à l'Algèbre abstraite* dans L. Boi /D. Flament/J.-M. Salanskis (ed.), 1830-1930 : *A Century of Geometry. Epistemology, History and Mathematics*, Berlin/Heidelberg/New York : Springer, 167-174, 1992.

[17] H. Sinaceur, *La dialectique de l'espace* dans M. Panza/J.-C. Pont, *Espace et horizon de réalité. Philosophie mathématique de Ferdinand Gonseth*, Paris/Milan/Barcelone/Bonn : Masson, 67-82, 1992.

Espace et temps physiques et description des systèmes mécaniques

Charles-Michel Marle[1]

Institut de Mathématiques
Université Pierre et Marie Curie
4 place Jussieu
75252 Paris cedex 05, France

1 Introduction

Les concepts d'espace et de temps se forment dans notre esprit à partir de la conscience que nous avons de l'Univers dans lequel nous vivons, grâce aux constructions intellectuelles que nous faisons pour ordonner nos expériences et nos perceptions, et tenter de "comprendre" cet Univers. Voici ce qu'Hermann Weyl (1885–1955) écrit à ce sujet, dans son beau livre "Temps, espace, matière" [32] :

> À propos du temps : *"La forme originelle du courant de conscience est le temps. C'est un fait qui peut être incompréhensible pour la raison, mais qui ne peut être nié : le contenu de la conscience n'est pas donné comme étant simplement, mais il est maintenant ; il remplit la forme du présent qui dure, par un contenu changeant ; de telle manière qu'il ne faut pas dire : cela est, mais cela est maintenant, ou cela n'est plus maintenant. Si par la réflexion on cherche à concevoir le contenu de ce courant de pensée comme un objet, il nous apparaîtra comme un écoulement dans le temps, dont les stades successifs se distinguent par les relations de plus tôt et plus tard".*
>
> À propos de l'espace : *"De même que le temps est la forme de ce courant de*

[1]marle@math.jussieu.fr

À la mémoire de mes maîtres, Alfred Senkeisen, Edgar Durupt, Marcel Saint-Jean et André Lichnerowicz

conscience, de même nous pouvons affirmer de plein droit que l'espace est la forme de la réalité corporelle. Les choses corporelles, telles que nous les percevons, possèdent en soi une extension spatiale. Mais c'est seulement quand l'Univers réel, unique et connexe, se construit avec nos expériences répétées, que l'extension spatiale donnée par chaque perception devient une partie d'un espace, toujours le même, qui embrasse toute chose. Cet espace est la forme du monde extérieur ; ce qui signifie que chaque chose corporelle peut occuper, sans être modifiée, une autre place que celle qu'elle occupe dans l'espace. Ce fait caractérise l'homogénéité de l'espace, il est la racine du concept de congruence".

Les philosophes ont abondamment discuté de la nature de l'espace et du temps. Le lecteur intéressé par les aspects historiques pourra consulter, par exemple, les ouvrages de B. C. van Frassen [29] et de M. Friedman [9], et trouvera dans les ouvrages de Poincaré [[25] (a), (b), (c), (d)] et de Weyl [32] de très intéressantes idées sur l'espace, le temps et la perception que nous en avons. Le grand philosophe Henri Bergson (1859–1941) a tenté, dans un petit livre [2], de concilier notre perception intuitive d'un temps absolu avec la théorie de la Relativité d'Einstein (cette tentative ne nous a pas semblé concluante). Mais pour moi du moins, la nature profonde de l'espace et du temps demeure tout à fait mystérieuse, tout comme restent mystérieuses la nature profonde de l'Univers et celle de la conscience humaine. Aussi me contenterai-je, dans ce qui suit, de traiter des *représentations mathématiques* de l'espace et du temps employées notamment en mécanique. Cette démarche, qui consiste à remplacer un objet qui nous est extérieur (l'espace, le temps, l'univers) par un concept mathématique, c'est-à-dire par une création de notre esprit, nous donne souvent l'impression de mieux comprendre l'objet que nous étudions. Acceptons ce fait, puisque c'est ainsi que l'esprit humain fonctionne, tout en sachant que nos constructions intellectuelles ne se confondent pas avec les objets qu'elles veulent représenter.

2 Espace et temps absolus de la mécanique newtonienne

2.1 Temps absolu

Notre intuition sensible nous conduit à représenter le temps comme un "continu à une dimension", car nous pouvons, par la pensée, ranger les événements par ordre chronologique en les disposant le long d'une ligne.

Que faut-il entendre par "continu à une dimension" ? Les progrès des mathématiques permettent de donner à ce concept un sens précis : il faut entendre par cela une *variété topologique réelle connexe de dimension* 1. Nous noterons $\mathcal{T}$ la variété qui représente le temps, et par abus de langage nous l'appellerons *temps* ; les éléments de $\mathcal{T}$ seront appelés *instants*.

Afin de lui conférer les propriétés nécessaires à son utilisation, nous allons munir le temps $\mathcal{T}$ de structures un peu plus riches que celle de variété topologique.

Tout d'abord, comme nous savons distinguer (au moins localement) entre deux instants lequel est antérieur à l'autre, nous sommes conduits à munir $\mathcal{T}$ d'une *orientation*, c'est-à-dire d'un sens de parcours privilégié, allant du passé vers le futur.

De plus, depuis l'invention du calcul différentiel par Newton et Leibniz, il est souhai-

table de pouvoir représenter l'évolution de systèmes physiques au cours du temps au moyen d'équations différentielles. Cela nous conduit à munir $\mathcal{T}$ de la structure mathématique adéquate, celle de *variété différentiable*.

Enfin, l'existence d'horloges (les systèmes planétaires, les rayonnements émis par des atomes excités, ...) permettent de comparer entre eux des intervalles de temps, voire même de les mesurer (en les comparant à un intervalle de temps choisi comme unité). Cela nous conduit à munir le temps d'une structure encore plus précise, celle de *variété affine*. Voici comment une telle structure permet la comparaison d'intervalles de temps. Pour tout couple d'instants t_1 et t_2 éléments de $\mathcal{T}$, l'intervalle de temps $t_2 - t_1$ est élément d'un espace vectoriel réel de dimension 1, noté $\overrightarrow{\mathcal{T}}$, appelé *espace des durées*. L'espace $\overrightarrow{\mathcal{T}}$ opère sur $\mathcal{T}$ (éventuellement localement seulement, si la variété affine $\mathcal{T}$ n'est pas complète) ; étant donné un troisième instant $t_3 \in \mathcal{T}$, cette opération permet de déterminer l'instant $t_4 = t_3 + (t_2 - t_1)$, tel que les intervalles de temps $t_4 - t_3$ et $t_2 - t_1$ soient égaux.

Moyennant le choix d'une durée positive particulière prise comme unité, l'espace vectoriel $\overrightarrow{\mathcal{T}}$ des durées peut être identifié à la droite réelle $\mathfrak{R}$. Mais ce choix d'une unité de temps n'a rien de canonique : la description mathématique des phénomènes physiques doit pouvoir être faite indépendamment de tout choix d'unités.

On connaît toutes les variétés affines réelles connexes orientées de dimension 1. À un isomorphisme affine près, ce sont, si l'on ne considère que les variétés sans bord :

- - la droite affine réelle orientée entière $\mathfrak{R}$,
- - les demi-droites affines ouvertes $]-\infty, 0[$ et $]0, +\infty[$,
- - l'intervalle ouvert borné $]0, 1[$,
- - le cercle trigonométrique orienté S^1.

Tous ces modèles sont localement isomorphes. C'est pourquoi lorsqu'on s'intéresse à l'évolution locale d'un système physique pendant un intervalle de temps limité, il n'est pas nécessaire de choisir parmi eux. Par contre, ce choix a son importance lorsqu'on traite de problèmes cosmologiques faisant intervenir le temps dans sa totalité. On devrait alors considérer comme possibles aussi les variétés à bord, c'est-à-dire ajouter à la liste ci-dessus :

- - les demi-droites affines fermées $]-\infty, 0]$ et $[0, +\infty[$,
- - les intervalles semi-ouverts bornés $[0, 1[$ et $]0, 1]$,
- - l'intervalle fermé borné $[0, 1]$.

Bien entendu, le choix entre ces divers modèles du temps n'a qu'un intérêt historique et anecdotique, car on sait maintenant que les problèmes cosmologiques ne peuvent être correctement traités dans le cadre de la mécanique classique.

Pour Isaac Newton (1642–1727) [20], le modèle qu'il convient d'adopter pour le temps est la droite affine réelle orientée entière $\mathfrak{R}$. C'est ce modèle qui est le plus souvent employé en Mécanique classique non relativiste. Signalons toutefois que selon certains philosophes, par exemple Friedrich Nietzsche (1844–1900) [21], le temps serait cyclique (théorie de l'*éternel retour*) ; pour eux, le temps serait donc isomorphe au cercle trigonométrique orienté S^1. Cette théorie, invérifiable par sa nature même, doit à notre avis être considérée comme un essai poétique plutôt que comme une théorie physique.

Nous devons maintenant évoquer l'importante question du caractère *universel* du temps. Assez curieusement, ce caractère a été implicitement admis sans discussion jusqu'à l'avènement de la théorie de la Relativité restreinte. En Mécanique newtonienne classique, on admet donc

qu'il existe un temps absolu $\mathcal{T}$, et qu'à tout événement on peut associer, de manière canonique (et non pas selon une convention ayant une part d'arbitraire), quel que soit l'endroit de l'espace où se produit cet événement, un élément unique de $\mathcal{T}$, qui est l'*instant* auquel a lieu cet événement.

2.2 Espace absolu

Notre intuition sensible nous conduit à représenter l'espace comme un "continu à trois dimensions", c'est-à-dire comme une variété topologique réelle connexe de dimension 3. L'existence de corps rigides utilisables comme règles et permettant de comparer entre elles des longueurs, la possibilité de faire des visées, de mesurer des angles, le fait que les rayons lumineux nous donnent l'intuition de la ligne droite, ont conduit à munir l'espace de structures bien plus riches que celle de variété topologique. Très tôt (les *Éléments* d'Euclide datent de 300 ans avant Jésus-Christ) il a été admis que l'espace pouvait être représenté mathématiquement par une partie d'un espace affine réel de dimension 3, et que dans cet espace certaines transformations étaient privilégiées. Ces transformations, appelées *déplacements euclidiens*, forment un groupe, le *groupe des déplacements euclidiens* ; il est engendré par les translations et par les rotations autour d'un point. Cette découverte est sans doute due à l'existence de corps solides que l'on peut déplacer dans l'espace, mettre dans diverses positions, et utiliser pour repérer la position d'autres objets par rapport à ces solides particuliers utilisés comme repères.

On dit souvent que l'espace physique est représenté mathématiquement par un espace affine réel $\mathcal{E}$, de dimension 3, muni d'une structure euclidienne (structure permettant de définir la *distance* $|\overrightarrow{AB}|$ de deux points A et B de $\mathcal{E}$, ainsi que le *produit scalaire* $g(\overrightarrow{V}_1, \overrightarrow{V}_2)$ de deux vecteurs $\overrightarrow{V}_1$ et $\overrightarrow{V}_2$, éléments de l'espace vectoriel $\overrightarrow{\mathcal{E}}$ associé à l'espace affine $\mathcal{E}$, cette distance et ce produit scalaire étant des fonctions à valeurs réelles). Cela n'est pas tout à fait exact, car ce n'est qu'après le choix d'une unité de longueur que l'espace physique se trouve muni d'une structure euclidienne. Il est pourtant possible de définir, de manière tout à fait canonique, un espace vectoriel réel orienté $\overrightarrow{\mathcal{L}}$, de dimension 1, appelé *espace des distances* dans $\mathcal{E}$, et un autre espace vectoriel réel orienté $\overrightarrow{\mathcal{L}}^2$, de dimension 1 lui aussi, appelé *espace des produits euclidiens* dans $\mathcal{E}$ (en fait $\overrightarrow{\mathcal{L}}^2$ est tout simplement le produit tensoriel $\overrightarrow{\mathcal{L}} \otimes \overrightarrow{\mathcal{L}}$ de deux exemplaires de $\overrightarrow{\mathcal{L}}$) ; moyennant quoi la *distance* de deux points A et B de $\mathcal{E}$, et le *produit euclidien* (nous préférons cette dénomination à celle de produit scalaire) $g(\overrightarrow{V}_1, \overrightarrow{V}_2)$ de deux vecteurs $\overrightarrow{V}_1$ et $\overrightarrow{V}_2$, éléments de $\overrightarrow{\mathcal{E}}$, peuvent être canoniquement définis, sans qu'il soit nécessaire de choisir d'unité de longueur ; ce sont des éléments, respectivement, des espaces $\overrightarrow{\mathcal{L}}$ et $\overrightarrow{\mathcal{L}}^2$. Pour cela, le lecteur pourra se reporter à l'Appendice 1 ci-après. Nous dirons donc que l'espace physique est représenté par un espace affine réel, de dimension 3, muni d'une *structure euclidienne intrinsèque*.

Cette description de l'espace physique n'a été remise en question qu'au cours du XIX-ème siècle, après la découverte, par Nikolai Ivanovich Lobachevski (1792–1856), János Bolyai (1802–1860) et Bernhard Riemann (1826–1866) des géométries non euclidiennes. Dans sa contribution à ce livre J.-P. Luminet présente les idées actuelles sur la géométrie de l'espace physique.

Abordons maintenant la délicate question du caractère absolu de l'espace, intimement liée à celle du caractère absolu du mouvement. Notre intuition sensible nous donne accès, non pas aux points de l'espace ou aux instants du temps, mais à des *événements*, c'est-à-dire à des couples formés d'un point d'espace (le point où a lieu l'événement) et d'un instant, élément du temps $\mathcal{T}$ (l'instant auquel a lieu l'événement). Mais comment comparer l'espace à deux instants différents ? En d'autres termes, considérons un objet matériel, par exemple un morceau de sparadrap, et observons-le à deux instants différents t_1 et t_2. Comment savoir si le point de l'espace $\mathcal{E}$ où est situé cet objet à l'instant t_1 est le même que celui où il est situé à l'instant t_2 ? Les philosophes et les savants, depuis Aristote jusqu'à Copernic non inclus, dans leur grande majorité (nous verrons plus loin qu'il y eut quelques brillantes exceptions), ont considéré la Terre comme immobile. Par suite, si l'objet matériel en question (notre morceau de sparadrap) était partie d'un rouleau soigneusement rangé, pendant tout l'intervalle de temps $[\,t_1, t_2\,]$, dans la réserve d'une pharmacie installée sur la terre ferme, ces savants auraient répondu que le point de l'espace occupé par cet objet aux instants t_1 et t_2 était le même. Si par contre ce morceau de sparadrap se trouvait collé au doigt d'un voyageur [10], ces mêmes savants auraient répondu que les points de l'espace qu'il occupait aux instants t_1 et t_2 étaient différents.

En formulant prudemment l'hypothèse du mouvement orbital des planètes, Terre comprise, autour du Soleil, Copernic a modifié cette manière de voir les choses. Il ne s'agissait pas d'une modification radicale, simplement d'un changement de repère. Pour Copernic, un repère fixe était, non plus un repère lié à la Terre, mais un repère ayant son origine au centre du Soleil, formé par trois axes pointant vers trois étoiles lointaines (les directions des étoiles lointaines étant à l'époque considérées comme fixes).

Cette conception a été retenue par Newton [20], pour qui l'espace, tout comme le temps, avait un caractère absolu. Pour lui, les notions de repos et de mouvement avaient donc aussi un caractère absolu.

3 Temps absolu, caractère relatif de tout mouvement

3.1 L'idée du caractère relatif de tout mouvement

Pour un mathématicien accoutumé à raisonner dans l'*espace-temps*, le caractère absolu du temps signifie l'existence d'une projection canonique de l'espace-temps sur un ensemble appelé *temps*, et le caractère absolu de l'espace l'existence d'une projection canonique de l'espace-temps sur un ensemble appelé *espace*. Ce sont donc deux notions de même nature. Curieusement, alors que le caractère absolu du temps a été admis sans discussion jusqu'à l'avènement de la théorie de la Relativité restreinte, le caractère absolu de l'espace a très tôt été mis en question.

Nicolas de Cues (1401–1464) pensait que tout, dans l'Univers, était en mouvement, y compris notre planète, et que la place de celle-ci n'était nullement privilégiée.

Nicolas Copernic (1473–1543), dont nous avons parlé ci-dessus, n'a pas mis en cause le caractère absolu du mouvement ou du repos ; mais pour lui, le repère fixe est lié, non plus à la Terre, mais au Soleil et aux étoiles lointaines ; pour lui, toutes les planètes de notre système tournent autour du Soleil. Son ouvrage *"De revolutionibus Orbium Coelestium"*, qu'il a

longtemps hésité à publier, parut en 1543, l'année même de sa mort. Il échappa ainsi aux rigueurs de l'Église qui, en 1616, a retiré son ouvrage de la circulation "jusqu'à correction" et en a interdit l'enseignement. Le système de Copernic s'est quand même imposé, grâce aux travaux de Galilée (1564–1642) (condamné en 1633 à abjurer publiquement sa découverte du mouvement de la Terre, et dont les livres furent brûlés) ainsi qu'à la découverte par Johannes Kepler (1571–1630) des trois très importantes lois portant son nom, qui régissent les mouvements des planètes du système solaire.

Giordano Bruno (1548–1600) imaginait un univers infini et, le premier, a compris de manière parfaitement claire le caractère relatif de tout mouvement. Afin de réfuter un argument employé pour nier le mouvement de la Terre, selon lequel si la Terre se déplaçait, une pierre lâchée du haut d'une tour ne tomberait pas verticalement au pied de cette tour, il a imaginé que la pierre était lâchée du haut du mât d'un navire en mouvement, et a écrit (d'après [26]) :

> *"Ainsi se meut avec la Terre tout ce qui se trouve sur la Terre (...) ; un homme placé sur le mât du navire, quelle que soit la vitesse de celui-ci, ne manquera pas le but ; rien n'empêchera la pierre ou tout autre objet pesant jeté (... du sommet) du mât (... d'atteindre) en ligne droite la base du mât".*

Son admirable indépendance d'esprit, son non conformisme intellectuel, lui ont valu un châtiment bien cruel : la Sainte Inquisition l'excommunia et le fit brûler vif. Il n'a toujours pas été réhabilité (alors que la révision du procès de 1633, engagée par le pape Jean Paul II, a abouti en 1992 à la réhabilitation de Galilée).

Wilhelm Gotfried Leibniz (1678–1716) a critiqué la conception que se faisait Newton d'un espace absolu, et suggéré qu'on exprime les lois de la mécanique sous une forme ne faisant intervenir que les positions relatives des corps matériels les uns par rapport aux autres, non leurs positions dans un hypothétique espace absolu. Malheureusement, une théorie mécanique qui ne fait intervenir que les distances mutuelles des corps matériels en présence ne permet de traiter correctement l'évolution d'un système isolé que si le moment cinétique total de ce système est nul. La suggestion de Leibniz ne pouvait donc pas aboutir avec les moyens connus à son époque.

Plus d'un siècle plus tard, Ernst Mach (1838–1916) [15] a repris, en les poussant plus loin encore, les idées de Leibniz. Pour lui, l'inertie d'un corps matériel n'existe qu'en raison des interactions de ce corps avec tous les autres corps matériels de l'Univers. Ses idées, difficiles à formaliser dans le cadre de la mécanique non relativiste, ont influencé Einstein dans sa conception de la Relativité générale[2].

Ces voix sont longtemps restées isolées. Cependant, les progrès de l'Astronomie et l'étude de la Dynamique ont conduit à remettre en question le caractère absolu du mouvement, donc de l'espace.

William Herschel (1738–1822) avait compris que la Voie Lactée est une galaxie, c'est-à-dire un amas d'étoiles en forme de disque, dont notre Soleil fait partie. Il pensait que le Soleil était proche de son centre. En étudiant le mouvement apparent de sept étoiles relativement proches, il fut le premier à voir qu'on devait attribuer un mouvement propre au Soleil. Les observations de Jacobus Cornelius Kapteyn (1851–1922) ont confirmé les idées

[2] voir [[8] (c)] pages 88 à 9

de Herschel sur l'appartenance du Soleil à une galaxie et sur le mouvement propre du Soleil. Le dogme de l'immobilité du Soleil se trouvait ainsi mis en doute. Harlow Shapley (1885–1972) a confirmé les résultats de Herschel et de Kapteyn sur la structure de la galaxie, tout en révisant fortement leurs estimations quant à ses dimensions et en montrant que la place du Soleil n'est pas du tout centrale. Bertil Lindblad (1895–1965) et Jan Hendrik Oort (1900–1992) ont découvert le mouvement de rotation d'ensemble de notre galaxie autour de son centre, la vitesse angulaire des étoiles dans ce mouvement étant d'autant plus grande que ces étoiles sont plus proches du centre de la galaxie. La nature extra-galactiques de certains objets, tels que la nébuleuse d'Andromède (M 31) a été reconnue par Herbert Doust Curtis (1872–1942). Edwin Powell Hubble (1889–1953) a observé la présence de céphéides dans la nébuleuse d'Andromède et conclu que cet objet était une autre galaxie, semblable à la nôtre. L'existence d'autres galaxies avait été pressentie par le philosophe Emmanuel Kant (1724–1804). De nombreuses galaxies ont été identifiées au cours du vingtième siècle. Vesto Melvin Slipher (1875–1969) a observé qu'Andromède se rapprochait de nous, à une vitesse d'environ 300 kilomètres par seconde. Avec Hubble , il a observé que la plupart des autres galaxies s'éloignent de nous. Milton La Salle Humason (1891–1973) et Hubble ont formulé la loi générale de fuite des galaxies : elles s'éloignent les unes des autres (et, en particulier, de la nôtre) à une vitesse proportionnelle à leurs distances mutuelles. Cette très importante découverte (dont l'interprétation est contestée par certains) est une pièce maîtresse des théories cosmologiques actuelles.

L'étude de la Dynamique a très tôt montré que les équations qui décrivent le mouvement relatif d'un système isolé par rapport à un repère qui, lui-même, est animé d'un mouvement de translation rectiligne et uniforme dans l'espace absolu, gardent la même forme que celles qui décrivent le mouvement de ce même système dans l'espace absolu. Voici ce que Paul Painlevé (1863–1933) écrit dans son petit livre, datant de 1909, sur les axiomes de la mécanique [23] :

> *"En un mot, le repérage absolu des mouvements n'est défini de par les axiomes de la Mécanique qu'à une translation près rectiligne et uniforme. Cette indétermination, sur laquelle nous revenons plus loin, ne présente d'ailleurs aucun inconvénient quant au développement de la Mécanique ; qu'on adopte un quelconque des repérages absolus entre lesquels on a ainsi le choix, on démontre que les forces absolues telles que nous les avons définies restent les mêmes".*

Dans ce petit livre, Painlevé revient en effet à plusieurs reprises sur cette indétermination ([23], pages 20, 27 à 31, 35 et 36), mais ne va pas jusqu'à mettre en doute l'existence de l'espace absolu. Par contre Poincaré [[25] (b), (c), (d)] rejette l'idée d'un espace absolu et affirme avec force le caractère relatif de tout mouvement. Dans les traités plus récents, tels que ceux de Jean Mandel [16], de Joseph Pérès [24] et de Henri Cabannes [22], tous les mouvements sont présentés comme relatifs, et les repères inertiels (relativement auxquels le mouvement d'un point matériel qui n'est soumis à aucune force est rectiligne et uniforme) sont considérés comme tous équivalents, aucun n'étant privilégié.

3.2 L'espace-temps fibré sur le temps et sa structure affine

Dans les traités classiques de mécanique [22, 16, 24], la présentation du caractère relatif de tout mouvement et l'introduction des repères inertiels est en général faite comme suit. L'espace est implicitement identifié à l'espace euclidien de dimension 3 de la géométrie. Cet espace est peuplé de corps, matériels ou conceptuels, avec parmi eux des solides qui peuvent être employés comme repères, dont la position dépend d'un paramètre, le temps. On évite d'évoquer la position absolue de chacun de ces corps dans l'espace, pour ne considérer que les positions relatives qu'ils occupent les uns par rapport aux autres. Chaque solide mobile peut servir à étudier le mouvement relatif des autres corps, par rapport à lui, utilisé comme repère. Le principe de l'inertie est alors introduit sous la forme suivante ([24], page 3) :

> *"Il existe au moins un repère géométrique, dit repère absolu, et un système de repérage du temps, (temps absolu) pour lesquels valent les énoncés ci-dessous :*
>
> *a) Un point matériel qui n'est soumis à aucune force a un mouvement d'accélération nulle, donc rectiligne et uniforme (principe de l'inertie) ;*
>
> *b) Un point matériel soumis à une ou plusieurs forces a, à chaque instant, une accélération proportionnelle à la résultante de ces forces ;*
>
> *c) Les coefficients de proportionnalité entre accélérations et résultante des forces sont des constantes caractéristiques des divers points matériels que l'on peut envisager. Ils sont, par définition, les masses".*

Il est alors facile de montrer que le repère appelé ci-dessus *repère absolu* (et qui, à notre avis, devrait plutôt être appelé *repère inertiel*) n'est pas unique, mais qu'il en existe une infinité, et que si l'on considère deux quelconques d'entre eux, le mouvement relatif de l'un par rapport à l'autre est un mouvement de translation uniforme.

Si elle est parfaitement cohérente et adaptée à l'étude des systèmes mécaniques rencontrés en pratique, cette présentation de la Mécanique classique a l'inconvénient de ne pas mettre clairement en évidence la structure géométrique de l'espace-temps : celle-ci n'apparaît qu'à travers les formules de changement de repères inertiels (ce qui peut convenir à un algébriste, mais ne satisfera pas un géomètre). On peut également lui reprocher de ne pas donner une description intrinsèque de l'évolution des systèmes mécaniques, puisque parmi les grandeurs qu'elle utilise, certaines (telles que les vitesses, les impulsions, l'énergie cinétique) ne dépendent pas que du système mécanique considéré, mais aussi du repère par rapport auquel on étudie son mouvement. Pour un géomètre, une description vraiment intrinsèque du mouvement serait beaucoup plus satisfaisante.

Élie Cartan (1869–1951), dans son mémoire [[4] (a)], a clairement décrit la structure géométrique de l'espace-temps de la Mécanique classique, et en a même proposé diverses généralisations. Ses réflexions sur ce sujet ont été très fécondes : il semble bien qu'elles aient contribué à lui faire découvrir diverses structures géométriques importantes, qu'il a appelées *connexions affines* [[4] (a)], *connexions conformes* et *connexions projectives* [[4] (b)]. Ces travaux (avec ceux de H. Weyl [32], eux aussi inspirés à ce grand savant par ses réflexions sur les théories physiques) sont à l'origine du concept de *connexion principale*, développé plus tard par Charles Ehresmann[3] (1905–1979) [7].

[3] Pour un exposé récent sur le sujet, voir l'article de Paulette Libermann [12].

Nous allons nous efforcer de donner un aperçu des idées de Cartan sur la structure géométrique de l'espace-temps en Mécanique non relativiste. Dans cette théorie, le temps $\mathcal{T}$ est absolu, mais l'espace ne l'est pas. Nous devons considérer qu'à chaque instant $t \in \mathcal{T}$, il existe un *espace à l'instant* t, noté $\mathcal{E}_t$, et que pour deux instants distincts t_1 et t_2, les espaces $\mathcal{E}_{t_1}$ et $\mathcal{E}_{t_2}$ sont disjoints. La réunion $\mathcal{U} = \bigcup_{t\in\mathcal{T}} \mathcal{E}_t$ des espaces à tous les instants est l'*espace-temps* ; ses éléments sont appelés *événements*. Il est muni d'une structure d'*espace fibré* dont la *base* est le temps $\mathcal{T}$, et dont la *projection canonique* est l'application qui, à chaque événement, associe l'unique instant t tel que cet événement soit élément de $\mathcal{E}_t$. Pour un instant $t \in \mathcal{T}$ donné, l'espace $\mathcal{E}_t$ à l'instant t est la *fibre* de l'espace-temps à cet instant. Chacune de ces fibres est munie, comme nous l'avons vu, d'une structure d'espace affine réel de dimension 3 et d'une structure euclidienne intrinsèque, au sens de l'Annexe 1. Mais nous devons encore munir l'espace-temps lui-même d'une structure géométrique permettant de comparer entre elles les fibres à différents instants. Comme nous l'avons vu ci-dessus, ce sont des propriétés mécaniques et physiques (l'existence de corps solides qu'on peut déplacer dans l'espace et employer comme règles, l'existence de rayons lumineux nous donnant l'intuition de la ligne droite) qui sont à l'origine de la création de la géométrie euclidienne. De même, c'est un principe de Mécanique, le *principe de l'inertie* énoncé ci-dessus, qui va nous permettre de munir l'espace-temps de la structure géométrique adéquate.

Considérons, pour chaque instant t, un repère affine $\big(P(t), \overrightarrow{e}_1(t), \overrightarrow{e}_2(t), \overrightarrow{e}_3(t)\big)$ de l'espace $\mathcal{E}_t$. Son origine $P(t)$ est un point de $\mathcal{E}_t$, et les vecteurs $\overrightarrow{e}_i(t)$ $(1 \leq i \leq 3)$ sont des éléments de l'espace vectoriel $\overrightarrow{\mathcal{E}}_t$ associé à l'espace affine $\mathcal{E}_t$ qui forment une base de cet espace. Nous supposons que cette famille de repères affines constitue un *repère inertiel* de l'espace-temps, qui pourrait être réalisé matériellement par un certain nombre de points matériels dont les distances relatives restent invariantes au cours du temps (comme les distances relatives des points d'un corps solide) et dont le mouvement est libre, aucune force ne s'exerçant sur ces points. Ayant ainsi construit, pour tout instant $t \in \mathcal{T}$, un repère affine de $\mathcal{E}_t$, nous pouvons identifier entre eux tous ces espaces, en identifiant, pour deux instants distincts t et t', le point de $\mathcal{E}_t$ dont les coordonnées, dans le repère $\big(P(t), \overrightarrow{e}_1(t), \overrightarrow{e}_2(t), \overrightarrow{e}_3(t)\big)$ sont (x_1, x_2, x_3) avec le point de $\mathcal{E}_{t'}$ dont les coordonnées, dans le repère $\big(P(t'), \overrightarrow{e}_1(t'), \overrightarrow{e}_2(t'), \overrightarrow{e}_3(t')\big)$ sont les mêmes, (x_1, x_2, x_3). Remarquons au passage que dans ces identifications, les espaces des distances $\overrightarrow{\mathcal{L}}_t$ et $\overrightarrow{\mathcal{L}}_{t'}$ dans $\mathcal{E}_t$ et $\mathcal{E}_{t'}$, ainsi que les espaces des produits euclidiens $\overrightarrow{\mathcal{L}}_t^2$ et $\overrightarrow{\mathcal{L}}_{t'}^2$ dans ces deux espaces, s'identifient, et que les structures euclidiennes intrinsèques des fibres $\mathcal{E}_t$ et $\mathcal{E}_{t'}$ se correspondent. Nous construisons ainsi une *trivialisation* de l'espace-temps $\mathcal{U}$. Comme la fibre-type $\mathcal{E}$ (c'est-à-dire l'espace auquel toutes les fibres $\mathcal{E}_t$ ont été identifiées) est un espace affine réel de dimension 3, et que la base $\mathcal{T}$ est un espace affine réel de dimension 1, l'espace-temps, identifié, grâce à la trivialisation ainsi construite, au produit $\mathcal{T} \times \mathcal{E}$, se trouve muni d'une structure d'espace affine réel de dimension 4. Pour tout couple (z_1, z_2) d'événements éléments de $\mathcal{U}$, nous pouvons donc donner un sens à la notion de vecteur $\overrightarrow{z_1 z_2}$ d'origine z_1 et d'extrémité z_2 (alors qu'auparavant nous ne pouvions le faire que si z_1 et z_2 étaient éléments d'une même fibre $\mathcal{E}_t$ en utilisant la structure d'espace affine de cette fibre). L'ensemble de ces vecteurs est l'espace vectoriel $\overrightarrow{\mathcal{U}}$ associé à l'espace affine $\mathcal{U}$, que nous pouvons maintenant faire opérer sur $\mathcal{U}$ par translations (ces translations pouvant désormais porter à la fois sur l'espace et sur le temps).

En utilisant le fait que le mouvement relatif d'un repère inertiel par rapport à un autre repère inertiel est une translation uniforme, il est facile de vérifier que la structure affine de $\mathcal{U}$ ainsi construite, donc aussi l'espace vectoriel $\overrightarrow{\mathcal{U}}$ et son action sur $\mathcal{U}$ par translations, ne dépendent pas du choix du repère inertiel qui a servi à trivialiser $\mathcal{U}$. Rejeter l'existence d'un espace absolu équivaut à considérer que parmi les trivialisations de $\mathcal{U}$ ainsi construites au moyen de repères inertiels, aucune n'est privilégiée. Cela n'empêche nullement la structure d'espace affine de $\mathcal{U}$ d'être parfaitement bien définie.

Cartan va plus loin dans son analyse, en remarquant que le rôle essentiel des repères inertiels est de permettre de définir le *transport parallèle* de vecteurs le long d'une courbe d'univers (courbe dans l'espace-temps), ce qui peut être fait au moyen d'une structure géométrique locale qu'il appelle *connexion affine* (et qu'on appelle aujourd'hui *connexion de Cartan*). Dans l'espace affine $\mathcal{U}$, l'équipollence de deux vecteurs, attachés à deux points distincts de $\mathcal{U}$, exprime le fait qu'ils se déduisent l'un de l'autre par transport parallèle le long d'une courbe joignant ces deux points. La connexion naturellement associée à la structure d'espace affine de $\mathcal{U}$ est dite *plate* parce que le résultat ne dépend pas du choix de la courbe le long de laquelle se fait le transport parallèle. Cartan montre aussi qu'en utilisant des connexions plus générales, non plates, pouvant comporter une *courbure* et une *torsion*, on peut munir l'espace-temps d'une structure géométrique tenant compte de l'existence d'un champ de forces (comme par exemple la pesanteur) et qu'on obtient ainsi une théorie qui, tout en restant classique (c'est-à-dire admettant la notion de temps absolu) ressemble beaucoup à la théorie de la Relativité générale, car elle incorpore certaines forces dans la géométrie de l'espace-temps.

3.3 Le groupe de Galilée

La structure géométrique de l'espace-temps est très clairement décrite par Vladimir Arnold, dès le premier chapitre de son ouvrage [1] ; cet auteur en fait, à juste titre, le point de départ de son exposé des principes de la mécanique classique. L'espace-temps $\mathcal{U}$ est muni d'une structure d'espace affine réel de dimension 4, d'une projection canonique $\mathcal{U} \to \mathcal{T}$, qui est une application affine, sur un espace affine réel de dimension 1 appelé *temps*, dont les fibres $\mathcal{E}_t$ (qui sont des sous-espaces affines réels de dimension 3, puisque la projection canonique est une application affine) sont toutes munies d'une structure euclidienne intrinsèque. De plus, toute translation dans l'espace affine $\mathcal{U}$ dont la composante temporelle $\theta \in \overrightarrow{\mathcal{T}}$ est non nulle applique chaque fibre $\mathcal{E}_t$ sur la fibre $\mathcal{E}_{t+\theta}$ par un isomorphisme d'espaces munis de structures euclidiennes intrinsèques, au sens de l'appendice 1. Le *groupe de Galilée* est le groupe des transformations de $\mathcal{U}$ qui préservent cette structure. Pour sa description explicite, le lecteur pourra se reporter à l'ouvrage d'Arnold déjà cité [1], ou au livre de Jean-Marie Souriau ([27], pages 138 à 140).

3.4 Descriptions intrinsèques de l'évolution d'un système mécanique

La connexion affine de l'espace-temps $\mathcal{U}$ permet de mettre les équations qui décrivent l'évolution d'un système mécanique sous une forme intrinsèque, n'utilisant aucun repère arbitraire auquel rapporter les mouvements, ni aucun choix particulier d'unités de temps, de longueur ou de masse. Nous allons traiter l'exemple très simple d'un point matériel

dans un champ de forces ; la méthode pourrait s'appliquer à des systèmes beaucoup plus compliqués.

La masse du point matériel est un élément d'un espace vectoriel de dimension 1, l'*espace des masses*, que nous noterons $\overrightarrow{\mu}$ (nous réservons la lettre $\mathcal{M}$ pour l'espace de Minkowski dont il sera question plus loin). Le mouvement du point matériel est décrit par une application $t \mapsto c(t)$ de l'ensemble $\mathcal{T}$ des temps dans l'espace-temps $\mathcal{U}$. Cette application doit bien sûr être une *section* de la projection canonique $\tau : \mathcal{U} \to \mathcal{T}$, ce qui veut dire qu'elle doit vérifier $\tau \circ c = id_{\mathcal{T}}$. La différentielle de c au point $t \in \mathcal{T}$ est élément de l'espace $\mathcal{L}(\overrightarrow{\mathcal{T}}, \overrightarrow{\mathcal{U}})$ des applications linéaires de $\overrightarrow{\mathcal{T}}$ dans $\overrightarrow{\mathcal{U}}$, qui s'identifie au produit tensoriel $\overrightarrow{\mathcal{T}}^* \otimes \overrightarrow{\mathcal{U}}$, où $\overrightarrow{\mathcal{T}}^*$ désigne le dual de $\overrightarrow{\mathcal{T}}$. La force f (qui peut éventuellement dépendre du point de l'espace-temps occupé par le point matériel considéré) prend ses valeurs dans l'espace vectoriel $\overrightarrow{\mu} \otimes \overrightarrow{\mathcal{T}}^* \otimes \overrightarrow{\mathcal{T}}^* \otimes \overrightarrow{\mathcal{U}}$. De plus, la projection de la force sur $\overrightarrow{\mu} \otimes \overrightarrow{\mathcal{T}}^* \otimes \overrightarrow{\mathcal{T}}^* \otimes \overrightarrow{\mathcal{T}}$ (définie à partir de la projection canonique $\tau : \mathcal{U} \to \mathcal{T}$) doit être identiquement nulle (la force n'accélère pas l'écoulement du temps). L'équation du mouvement du point matériel s'écrit

$$m \otimes \nabla_{Dc(t)}\big(Dc(t)\big) = f\big(c(t)\big) .$$

Dans cette équation, $\nabla_{Dc(t)}\big(Dc(t)\big)$ est la dérivée covariante, pour la connexion affine de l'espace-temps $\mathcal{U}$, du vecteur $Dc(t)$ relativement à lui-même.

En Appendice 3 nous montrons comment cette équation, explicitée dans un système de coordonnées locales correspondant à un repère tournant, conduit à exprimer la force centrifuge et la force de Coriolis (considérées, en Mécanique, comme des "forces fictives") au moyen des symboles de Christoffel de la connexion dans ce système de coordonnées.

Nous allons brièvement présenter quelques travaux, effectués dans la période 1930-2000, qui nous semblent avoir été inspirés à leurs auteurs par leurs réflexions sur la description des systèmes mécaniques. Sans avoir le caractère exceptionnellement novateur des travaux d'Einstein, qui ont révolutionné notre manière de voir l'univers, ces travaux nous semblent dignes d'être mentionnés.

André Lichnerowicz (1915–1998) [[13] (a), (b), (c)], étudiant l'expression intrinsèque des équations de la Mécanique, a été conduit à introduire des structures géométriques nouvelles dont l'importance ne cesse de croître, qu'il a appelées *variétés de Poisson* [[13] (d)] et *variétés de Jacobi* [[13] (e)]. Il a notamment montré que le cadre naturel, pour l'étude de la géométrie des transformations canoniques employées en Mécanique analytique, était une *variété canonique* [[13] (a), (f)], c'est-à-dire une variété de Poisson de dimension impaire dont les feuilles symplectiques sont de codimension 1, l'ensemble de ces feuilles s'identifiant à l'ensemble des temps $\mathcal{T}$. Les très importants travaux d'André Lichnerowicz sur les variétés de Poisson l'ont notamment conduit à l'étude des déformations des structures de Poisson, conduisant à un méthode de quantification. Les variétés de Poisson, qui ont donné lieu à un important travail d'Alan Weinstein [[31] (a)], sont toujours très activement étudiées.

L'approche très originale de Jean-Marie Souriau [27] pour l'étude des systèmes mécaniques utilise, comme point de départ, l'*espace d'évolution* $\mathcal{E}$ *du système*, variété de dimension impaire $2n + 1$ (qui, en Mécanique non relativiste, n'est autre que la variété canonique d'André Lichnerowicz). Cette variété (dont chaque élément décrit un état cinématique possible du système dans l'espace-temps) est munie d'une 2-forme fermée dont le noyau, de

dimension 1, détermine un feuilletage en courbes de l'espace d'évolution. Chacune de ces courbes représente un *mouvement* du système. L'ensemble des mouvements est une variété $\mathcal{R}$ (éventuellement non séparée), de dimension paire $2n$, appelée *variété des mouvements* du système.

Lorsque le système considéré relève de la Mécanique non relativiste, on peut assez facilement comprendre pourquoi $\mathcal{R}$ a bien une structure de variété. Dans ce cas, il existe une projection canonique $\tau : \mathcal{E} \to \mathcal{T}$ de l'espace d'évolution du système sur l'ensemble des temps qui, à chaque état cinématique du système dans l'espace-temps, associe l'instant qui lui correspond. Considérons une valeur particulière t_0 du temps, et notons $\mathcal{E}_{t_0} = \tau^{-1}(t_0)$ la sous-variété de $\mathcal{E}$, de dimension paire $2n$, formée par les éléments de l'espace d'évolution dont la coordonnée "temps" est t_0. L'ensemble des mouvements qui existent à l'instant t_0 est un ouvert de l'ensemble de tous les mouvements, et l'application qui associe, à chaque élément de cet ouvert, le point unique où ce mouvement (qui est une courbe dans $\mathcal{E}$) rencontre $\mathcal{E}_{t_0}$ est une bijection de cet ouvert sur $\mathcal{E}_{t_0}$. On peut employer cette bijection pour construire une carte de $\mathcal{R}$ dont le domaine est l'ensemble des mouvements qui existent à l'instant t_0. Si on considère un autre instant t_1, on peut de même construire une autre carte de $\mathcal{R}$ dont le domaine est l'ensemble des mouvements qui existent à l'instant t_1. Le théorème de différentiabilité du flot d'une équation différentielle permet alors de voir que l'application de changement de carte est un difféomorphisme. C'est pourquoi l'ensemble des cartes de $\mathcal{R}$ qu'on peut ainsi construire, en considérant tous les instants possibles, détermine sur $\mathcal{R}$ une structure de variété différentiable. La variété des mouvements peut ne pas être séparée en raison de l'existence de mouvements non éternels, qui cessent d'exister (ou qui commencent à exister) pour une valeur finie du temps (comme, par exemple, les mouvements de collision ou d'éjection dans le problème des deux corps).

La 2-forme fermée dont l'espace d'évolution est muni détermine, par projection, une forme symplectique sur la variété des mouvements. L'idée de la variété des mouvements, munie de sa forme symplectique, se trouve en germe dans la *Mécanique analytique* de Joseph Louis Lagrange (1736–1813) [11], plus précisément dans son étude des équations qui régissent l'évolution des éléments orbitaux des planètes. Les symétries de l'espace d'évolution (notamment les translations dans le temps si le système est autonome, éventuellement aussi les transformations de Galilée) donnent, après projection, des symétries de la variété des mouvements, exprimées mathématiquement par l'action d'un groupe de Lie sur cette variété, cette action conservant la forme symplectique. Lorsque le système mécanique considéré est traité dans le cadre de la Mécanique classique, ce groupe de symétries contient le groupe de Galilée. Mais l'approche de Souriau reste encore valable en Mécanique relativiste, le groupe de symétries du système contenant alors le groupe de Poincaré (dont il sera question plus loin). Les variétés des mouvements sur lesquelles le groupe de Galilée dans le cas classique, ou le groupe de Poincaré dans le cas relativiste, agit de manière transitive, ont été classifiées par J.-M. Souriau (qui, au passage, a le premier obtenu un intéressant théorème sur la classification de tous les espaces homogènes symplectiques, obtenu indépendamment par Kirillov et Kostant). Cette classification a une interprétation physique : elle correspond à la classification des particules élémentaires selon leur *masse* et leur *spin*.

L'étude des systèmes mécaniques a conduit Wlodzimierz Tulczyjew [28] à de très intéressants développements concernant la géométrie, sur une variété différentiable M, des fibrés tan-

gents et cotangents itérés TT^*M, T^*TM et T^*T^*M. Ces outils mathématiques nouveaux lui ont permis de renouveler les formulations lagrangienne et hamiltonienne de la Mécanique.

Les importants travaux d'Alan Weinstein [[31] (b)] sur les groupoïdes symplectiques, puis sur les groupoïdes de Poisson, ont aussi pour origine, au moins en partie, les problèmes rencontrés en Mécanique.

4 Abandon du temps absolu : la Relativité restreinte

4.1 Vitesse de la lumière, expériences de Michelson et Morley

Dès 1675, l'astronome Ole Christensen Römer (1644–1710) avait compris que la lumière se propage à une vitesse finie et donné une estimation relativement précise de cette vitesse. Ce résultat, qu'il avait déduit de l'observation des instants auxquels il avait observé les éclipses des satellites de Jupiter, ne fut pleinement reconnu qu'après l'observation, en 1727, par l'astronome James Bradley (1693-1762), de l'aberration des étoiles lointaines. Ce phénomène (une légère variation de la direction apparente d'une étoile, lors d'observations faites à deux époques différentes de l'année) avait fourni à Bradley à la fois la première preuve directe du mouvement relatif de la Terre par rapport aux étoiles lointaines et une mesure de la vitesse de la lumière plus précise que l'évaluation faite par Römer.

Par ailleurs, James Clerk Maxwell (1831-1879), qui avait, vers 1873, établi les équations générales qui régissent l'évolution d'un champ électromagnétique, avait compris que la lumière est une onde électromagnétique. Ce fait fut très généralement admis après que l'on eut vérifié que la vitesse à laquelle elle se propage est bien la même que celle à laquelle se propagent les ondes hertziennes.

Or les équations de Maxwell , qui prévoient la propagation de la lumière à la même vitesse c dans toutes les directions, ne sont pas invariantes par le groupe de Galilée. Dans un repère mobile à une vitesse v par rapport à l'hypothétique milieu (appelé *éther*) servant de support aux ondes électromagnétiques, la vitesse de la lumière devrait être $c - v$ dans la direction de la vitesse v, $c + v$ dans la direction opposée, et devrait prendre, dans les directions transverses à la vitesse v, toutes les valeurs comprises entre $c - v$ et $c + v$, suivant leurs orientations. L'électromagnétisme faisait ainsi réapparaître l'hypothétique repère absolu, dont les mécaniciens avaient eu tant de mal à se débarrasser ! Des expériences, destinées à comparer la vitesse de la lumière dans deux directions orthogonales, furent réalisée par Albert Abraham Michelson (1852–1931) d'abord seul (1880) puis (en 1887 [17] et à plusieurs reprises par la suite) en collaboration avec Edward Williams Morley (1838–1923). Diverses hypothèses plus ou moins satisfaisantes furent proposées pour expliquer le résultat négatif de ces expériences. Par exemple, l'entraînement de l'éther par la Terre, envisagé par Michelson, s'accordait mal avec l'aberration des étoiles lointaines. Hendrik Anton Lorentz (1853–1928) proposa en 1895 [[14] (a)] une explication basée sur une contraction des longueurs d'un solide en mouvement selon la direction parallèle à sa vitesse. La même explication fut aussi proposée indépendamment par George Francis FitzGerald (1851–1901). En 1904 [[14] (b)], Lorentz observa l'invariance de la forme des équations de Maxwell lors de changements de variables (aujourd'hui appelés *transformations de Lorentz*) portant à la fois sur les coordonnées d'espace et le temps. Ainsi que l'a signalé Poincaré dans son mémoire [[25] (e)],

cette transformation avait été précédemment considérée par Woldemar Voigt (1850–1919) [30].

En 1905, presque simultanément et sans doute indépendamment, Albert Einstein (1879–1955) et Henri Poincaré (1854–1912) soumettent pour publication deux importants mémoires qui, tous deux, contiennent l'essentiel des bases mathématiques de la théorie de la Relativité restreinte.

Le mémoire d'Einstein [[8] (a)] met très clairement en cause la notion de temps absolu, en présentant une critique détaillée de la notion de simultanéité. Il propose d'admettre deux postulats fondamentaux :

- (i) aucun phénomène physique, qu'il soit mécanique ou électromagnétique, ne peut permettre de distinguer, parmi tous les repères inertiels, un éventuel repère absolu ;
- (ii) la vitesse de la lumière est la même dans toutes les directions, et ne dépend pas du repère inertiel par rapport auquel on l'évalue.

En utilisant ces deux postulats, Einstein retrouve les formules de changement de repère précédemment obtenues par Lorentz, discute des conséquences de la nouvelle théorie pour la cinématique (lois de composition des vitesses) et l'électromagnétisme (expressions d'un même champ électromagnétique dans deux repères en mouvement relatif l'un par rapport à l'autre). Il revient sur la contraction des longueurs postulée par Lorentz et sur le phénomène corrélatif de dilatation des temps, mais avec une explication beaucoup plus satisfaisante basée sur le caractère non absolu de la simultanéité. Enfin il observe l'équivalence fondamentale entre masse et énergie, donnée par la fameuse relation

$$E = mc^2 \,,$$

sur laquelle il reviendra lors du développement de la théorie de la relativité générale. Ces idées révolutionnaires sont explicitées et popularisées dans ses conférences et publications ultérieures [[8] (b), (c), (d)].

Le très riche mémoire de Poincaré [[25] (e)], certainement rédigé tout à fait indépendamment de celui d'Einstein, contient une étude exhaustive des transformations de Lorentz, considérées comme formant un groupe et ainsi rattachées à une théorie mathématique bien établie. Poincaré admet le premier postulat d'Einstein (impossibilité de distinguer entre repères inertiels), qu'il formule dans des termes très semblables à ceux employés par Einstein. Comme Einstein, il obtient les formules relativistes de composition des vitesses. Il discute de la dynamique de l'électron, observe qu'on doit distinguer sa masse longitudinale et sa masse transversale, et envisage une modification de la loi d'attraction universelle de Newton afin de la rendre compatible avec une transmission des interactions gravitationnelles à la vitesse de la lumière, qu'il considère comme une limite infranchissable. Comme Einstein, Poincaré a compris l'équivalence de la masse et de l'énergie, puisqu'il a cherché à prouver que la masse de l'électron était d'origine électromagnétique (explication aujourd'hui considérée comme inexacte).

Peut-on attribuer conjointement à Einstein et à Poincaré la découverte de la théorie de la Relativité restreinte ? Certains physiciens et historiens des sciences le pensent. Cependant, les idées physiques (caractère non absolu de la simultanéité, nécessité de raisonner sur les événements, dans l'espace-temps, plutôt que séparément sur les points d'espace et le temps) sont, à notre avis, beaucoup plus clairement exprimées par Einstein que par Poincaré.

Pour Poincaré, les transformations de Lorentz et les "temps locaux" qu'elles font intervenir semblent rester des artifices mathématiques. Il ne met pas en doute l'existence d'un temps absolu. Ayant fait d'admirables travaux dans le domaine de la Mécanique classique [[25] (f)], peut-être Poincaré avait-il quelque réticence à en remettre les bases en question. Peut-être aussi était-il trop imprégné par l'idée, à plusieurs reprises affirmée dans ses livres de Philosophie des sciences [[25] (a), (b), (c), (d)], selon laquelle une description du monde physique peut être plus commode qu'une autre sans pour autant être plus vraie, et qu'entre plusieurs théories rendant aussi bien compte des phénomènes physiques, il convient d'employer la plus simple, jusqu'à ce qu'une autre la remplace parce qu'elle se révèle encore plus commode et plus simple. Poincaré n'était pas convaincu de la plus grande simplicité ou de la plus grande commodité de la théorie de la Relativité ; voici en effet la conclusion d'une conférence qu'il a faite à Londres le 4 mai 1912, publiée dans ses *Dernières pensées* [[25] (d)] , dans laquelle il présentait cette théorie :

> *"Quelle va être notre position en face de ces nouvelles conceptions ? Allons-nous être forcés de modifier nos conclusions ? Non certes : nous avions adopté une convention parce qu'elle nous semblait commode, et nous disions que rien ne pourrait nous contraindre à l'abandonner. Aujourd'hui certains physiciens veulent adopter une convention nouvelle. Ce n'est pas qu'ils y soient contraints ; ils jugent cette convention nouvelle plus commode, voilà tout ; et ceux qui ne sont pas de cet avis peuvent légitimement conserver l'ancienne pour ne pas troubler leurs vieilles habitudes. Je crois, entre nous, que c'est ce qu'ils feront encore longtemps".*

Sur ce dernier point, Poincaré ne se trompait malheureusement pas ! Il est à noter aussi que dans ses conférences et écrits nettement postérieurs à la publication du mémoire d'Einstein, par exemple la conférence au congrès de Lille de 1909 et celle de 1912 dont la conclusion est citée ci-dessus, Poincaré expose les principes de la Relativité restreinte en rendant un hommage appuyé à Lorentz, mais sans jamais mentionner Einstein ! Il est vrai que de même Einstein, de son côté, ne mentionne jamais dans ses écrits le mémoire de Poincaré de 1905.

4.2 L'espace-temps de Minkowski

Comme en mécanique classique non relativiste, l'espace-temps de la Relativité restreinte est un espace affine réel $\mathcal{M}$, de dimension 4. Mais cet espace n'est plus doté d'une projection canonique sur un ensemble des temps, car il n'existe plus de temps absolu. Par contre, l'espace vectoriel $\overrightarrow{\mathcal{M}}$ associé à l'espace affine $\mathcal{M}$ est muni de ce que nous avons appelé une *structure de Minkowski intrinsèque* (voir Appendice 3), définie par une application bilinéaire symétrique de $\overrightarrow{\mathcal{M}} \times \overrightarrow{\mathcal{M}}$ dans un certain espace vectoriel réel de dimension 1, noté $\overrightarrow{\mathcal{Q}}^*$ (qui est le dual de l'espace vectoriel $\overrightarrow{\mathcal{Q}}$ engendré par une certaine forme quadratique de signature $(+,-,-,-)$ sur $\overrightarrow{\mathcal{M}}$). Nous montrons dans l'Appendice 3 que le choix d'un élément non nul particulier de $\overrightarrow{\mathcal{Q}}$ permet de définir, dans $\mathcal{M}$, à la fois une unité de longueur (utilisable pour la mesure de segments de droite sur les droites de genre espace) et une unité de temps (utilisable pour la mesure de segments de droite sur les droites de genre temps).

L'espace-temps $\mathcal{M}$ de la Relativité restreinte est nommé *espace de Minkowski*, en hommage à Hermann Minkowski (1864–1909) qui en a décrit la géométrie [18] et a eu, le premier,

l'idée d'utiliser une coordonnée imaginaire pure pour exprimer les transformations de Lorentz au moyen de formules analogues à celles qui expriment les rotations.

4.3 Le groupe de Poincaré

Le groupe de Lorentz, ainsi nommé par Poincaré, est le groupe des transformations linéaires de l'espace vectoriel $\overrightarrow{\mathcal{M}}$ qui préservent chaque forme quadratique élément de l'espace vectoriel $\overrightarrow{\mathcal{Q}}$, de dimension 1. Le groupe de Poincaré est le groupe des transformations de l'espace-temps de Minkowski $\mathcal{M}$ qui préservent sa structure. Il est engendré par les translations de cet espace affine et les transformations de Lorentz (l'espace affine $\mathcal{M}$ étant identifié à l'espace vectoriel associé en choisissant un point quelconque comme origine). Le lecteur en trouvera une description explicite dans le livre de J.-M. Souriau [27], page 170.

5 La Relativité générale

5.1 Insuffisances de la Relativité restreinte

Dans sa troisième conférence à l'université de Princeton [[8] (c)], Einstein indique très clairement les considérations qui l'ont conduit à développer la théorie de la Relativité générale. En voici un bref aperçu.

Dans la théorie de la Relativité restreinte, tout comme en Mécanique classique non relativiste, il existe une classe particulière de repères privilégiés : les repères inertiels. Einstein a cherché à construire une théorie dans laquelle aucun repère ne serait privilégié. D'ailleurs, le concept même de repère global, inspiré par l'existence de corps solides indéformables par rapport auxquels on peut repérer la position, en fonction du temps, des objets physiques, a-t-il encore un sens en Relativité ? On peut légitimement se poser cette question, puisque l'impossibilité de transmettre des signaux à distance à une vitesse supérieure à celle de la lumière est incompatible avec l'existence de corps solides parfaitement rigides. En étudiant la géométrie d'un disque tournant, Einstein a même été conduit à mettre en doute la validité de la géométrie euclidienne pour décrire les propriétés physiques des hyperplans de genre espace dans l'espace-temps.

D'autre part, Einstein remarque que la théorie de la Relativité restreinte n'explique pas un fait troublant : l'égalité de la masse inerte et de la masse gravitationnelle. Les expériences faites par Loránd Eötvös (1848–1919) avaient montré cette égalité (ou plutôt la constance du rapport de la masse gravitationnelle à la masse inerte pour des substances de différentes compositions chimiques), avec une précision de l'ordre de 5.10^{-9}. Des expériences faites par Robert Henry Dicke vers 1960 ont porté cette précision à 10^{-11}. Ce résultat implique (et équivaut) à l'équivalence entre champ gravitationnel et champ d'accélération : en rapportant les mouvements à un repère plus ou moins accéléré on peut, à volonté, faire apparaître ou annuler un champ gravitationnel dans une région restreinte de l'espace-temps. C'est ainsi que dans un avion en vol parabolique, simulant la chute libre, on peut faire disparaître la force de pesanteur. Einstein propose d'accorder à ce fait expérimental le rang de principe, et l'appelle *principe d'équivalence*. Pour en tenir compte, il a l'intuition géniale de traduire le champ gravitationnel en termes de propriétés géométriques de l'espace-temps.

Enfin, l'équivalence masse - énergie, déjà découverte lors du développement de la Relativité restreinte, conduit Einstein à penser que comme la masse, toute forme d'énergie (par exemple l'énergie électromagnétique) doit contribuer à la création d'un champ gravitationnel. Il y aurait ainsi réciprocité, comme dans la théorie de l'électromagnétisme de Maxwell : la masse d'un objet matériel, et plus généralement toute forme d'énergie, recevrait l'influence du champ de gravitation, ce qui affecterait son mouvement, et réciproquement contribuerait à créer ce champ, intervenant dans la structure géométrique de l'espace-temps.

Alors que la paternité de la théorie de la Relativité restreinte peut sembler à certains partagée entre Lorentz, Poincaré et Einstein, c'est sans conteste possible Einstein seul qui est à l'origine de la théorie de la Relativité générale.

5.2 Espace-temps et équations d'Einstein

Einstein conserve la structure locale de l'espace de Minkowski, c'est-à-dire la forme bilinéaire, de signature $(+,-,-,-)$, servant à mesurer à la fois les distances et les durées. Mail il lui confère un caractère purement local : elle dépend du point de l'espace-temps considéré. Ainsi l'espace-temps d'Einstein est une variété différentiable réelle, de dimension 4, munie de ce que l'on appelle aujourd'hui une métrique pseudo-riemannienne de signature $(+,-,-,-)$, notée g.

Comme en mécanique classique (paragraphe 2.2 et appendice 1) ou en Relativité restreinte (appendice 3), on peut remarquer que ce n'est qu'après le choix d'une unité de mesure que la métrique pseudo-riemannienne g est déterminée. En changeant d'unité, on remplace g par λg, où λ est un scalaire réel strictement positif. En procédant comme dans l'appendice 3, on peut s'affranchir complètement du choix d'une unité en remplaçant la métrique pseudo-riemannienne g par une application bilinéaire symétrique à valeurs dans l'espace vectoriel $\overrightarrow{\mathcal{Q}}^*$, de dimension 1, dual de l'espace $\overrightarrow{\mathcal{Q}}$ engendré par g.

La ligne d'univers d'une particule-test (soumise au champ gravitationnel, mais supposée de masse suffisamment faible pour ne pas contribuer de manière significative à sa création) est une géodésique de genre temps de la connexion de Levi-Civita associée à la métrique pseudo-riemannienne g.

Il restait à Einstein à formuler les équations qui indiquent comment la masse des objets matériels, et les diverses formes d'énergie, distribuées dans l'espace-temps, contribuent à la formation du champ de gravitation. Einstein semble s'être inspiré de l'équation de Poisson

$$\Delta\phi = 4\pi\rho$$

qui lie le champ d'attraction newtonienne ϕ créé par des masses distribuées dans l'espace, à la densité ρ de cette distribution de masses. Il pense à représenter la distribution d'énergie dans l'espace-temps par un champ T de tenseurs symétriques du second ordre appelé *tenseur d'impulsion-énergie*. Einstein aboutit à la célèbre équation

$$\mathcal{R} - \frac{1}{2}Rg = \chi T\,,$$

où $\mathcal{R}$ est le tenseur de Ricci de la métrique pseudo-riemannienne g et R la courbure scalaire (obtenue par contraction de $\mathcal{R}$). Au membre de droite, χ est une constante scalaire, et T le

tenseur d'impulsion-énergie. Le membre de gauche de l'équation ci-dessus est souvent appelé *tenseur d'Einstein* et noté $\mathcal{S}$.

Plus tard [[8] (d)], des considérations sur la structure globale de l'espace-temps (problème cosmologique), et le désir de rendre possible l'existence de solutions stationnaires (Einstein acceptait mal l'idée d'un univers en évolution) conduisent Einstein à modifier cette équation en ajoutant un terme :

$$\left(\mathcal{R} - \frac{1}{2}Rg\right) + \Lambda g = \chi T\,,$$

où Λ est une nouvelle constante, appelée *constante cosmologique*.

Dès 1911, Einstein avait fait une prévision pouvant servir de test de la théorie de la Relativité gènérale : un rayon lumineux, provenant d'une étoile lointaine et rasant le bord du Soleil, devait, selon cette théorie, subir une déflexion de 1,7 seconde. Cette prévision a été vérifiée par Arthur Stanley Eddington (1882–1944) lors de l'éclipse totale de 1919. Une autre observation astronomique (l'avance du périhélie de Mercure de 43 secondes par siècle en plus de l'avance prédite par la théorie des perturbations classique), a également très tôt étayé cette théorie. Un troisième effet (la déviation vers le rouge de la longueur d'onde de la lumière lorsqu'elle se déplace à contre-courant dans un champ gravitationnel) a été confirmé par des observations astronomiques en 1925 (déviation vers le rouge des raies spectrales des étoiles massives) et par une expérience terrestre faite par R. Pound et G. Rebka en 1959. Tout récemment, les prévisions de la théorie de la Relativité générale ont dû être appliquées à un problème technique : la détermination de la position des objets terrestres au moyen des ondes émises par des satellites artificiels ("Global Positioning System") ; les corrections relativistes sont en effet indispensables pour donner à cette détermination toute sa précision.

Après avoir créé la théorie de la Relativité générale, Einstein a cherché à construire une théorie plus générale encore qui ferait entrer le champ électromagnétique dans la géométrie de l'espace-temps. Ces recherches ont donné lieu à des échanges de lettres entre Einstein et Cartan [5] portant sur la théorie des connexions et le parallélisme absolu. La lecture de ces lettres est extrêmement enrichissante et instructive. La nouvelle théorie élaborée par Einstein, dans laquelle la connexion de Levi-Civita est remplacée par une connexion non symétrique, est brièvement présentée dans [[8] (e)] et discutée par Cartan [[4] (b)], pages 1863–1875. Cette théorie n'a malheureusement pas pleinement abouti.

6 Remerciements et conclusion

Je remercie les organisateurs du Colloque international "Géométrie au vingtième siècle, 1930–2000" de m'avoir invité à présenter ces quelques idées, bien qu'une bonne partie des travaux dont j'ai parlé soit quelque peu antérieure à 1930 (mais le temps n'est-il pas relatif ?). Je remercie également Alain Albouy pour d'intéressants échanges de vue sur l'usage d'unités en Mécanique et en Physique, Jean Gunther pour d'importantes précisions sur l'histoire de l'Astronomie, José Beltrán, Paulette Libermann et Roland Pons pour leur lecture attentive d'une version préliminaire de ce travail et leurs critiques constructives. Je remercie Alan Weinstein et l'expert anonyme qui a examiné ce travail pour leurs recommandations, dont j'ai essayé de tenir compte.

Les limites de mes compétences m'ont contraint à laisser de côté plusieurs questions importantes qu'on peut légitimement se poser à propos de l'espace et du temps. Qu'en est-il de l'irréversibilité du temps ? Quels sont les liens qui unissent la causalité et le sens de l'écoulement du temps ? Que peut-on dire du temps cosmologique, et peut-on légitimement parler de l'âge de l'univers, s'il n'existe pas de temps absolu ? Comment l'espace et le temps entrent-ils dans la description des phénomènes quantiques ?

Pour terminer cet exposé, je voudrais souligner le fait que les descriptions du temps et de l'espace données ci-dessus admettent toutes la possibilité de donner un sens au concept d'événement ponctuel de l'espace-temps. Est-ce vraiment légitime ? Peut-on vraiment concevoir un événement de durée nulle, d'extension spatiale nulle, correspondant à un point mathématique de la variété espace-temps ? Cette question avait déjà été soulevée par Poincaré [[25] a]. Les variétés différentiables sont des objets mathématiques qui, en dernier ressort, sont fondés sur les propriétés de l'ensemble $\mathfrak{R}$ des nombres réels. Cet ensemble est une construction mathématique due, dans une large mesure, à Julius Dedekind (1831–1916). Il est tout à fait admirable et stupéfiant que cette construction se soit révélée adaptée à la description du monde réel, à des échelles allant de 10^{-15} mètres (dimension des particules élémentaires) à 10^{26} mètres (dimension estimée de l'univers). De nouveaux objets mathématiques, tels que la géométrie non commutative d'Alain Connes, permettront peut-être, dans l'avenir, d'aller encore plus loin dans la compréhension de l'univers.

7 Appendice

Appendice 1. Espaces des distances et des produits euclidiens

Soit $\mathcal{E}$ un espace affine réel, de dimension 3, pris comme modèle de l'espace physique, sur lequel agit le groupe G_E des déplacements euclidiens. Nous allons voir comment, de manière tout à fait canonique, sans avoir à choisir d'unité de longueur, munir cet espace d'une structure (que nous appellerons *structure euclidienne intrinsèque*) permettant de définir la distance de deux points de $\mathcal{E}$ et le produit euclidien de deux vecteurs éléments de l'espace vectoriel $\overrightarrow{\mathcal{E}}$ associé à $\mathcal{E}$.

Considérons l'ensemble de toutes les droites vectorielles orientées plongées dans $\mathcal{E}$ (c'est-à-dire l'ensemble de toutes les droites affines orientées de $\mathcal{E}$, chacune de ces droites étant munie d'un point privilégié jouant, sur cette droite, le rôle d'origine). Le groupe G_E des déplacements euclidiens opère transitivement sur cet ensemble. De plus, si $\overrightarrow{D}$ est une droite vectorielle orientée plongée dans $\mathcal{E}$, et g un élément du groupe G_E qui applique $\overrightarrow{D}$ sur elle-même par un isomorphisme vectoriel préservant l'orientation, la restriction de g à $\overrightarrow{D}$ est l'application identique. Par suite, le quotient de l'ensemble de toutes les droites vectorielles orientées plongées dans $\mathcal{E}$ par l'action du groupe G_E est un espace vectoriel orienté, de dimension 1, que nous appellerons *espace des distances* dans $\mathcal{E}$, et que nous noterons $\overrightarrow{\mathcal{L}}$.

Pour tout couple (A, B) de points de $\mathcal{E}$, nous pouvons définir la distance $|\overrightarrow{AB}|$ de ces deux points comme suit : c'est l'unique élément de l'espace $\overrightarrow{\mathcal{L}}$ des distances qui est l'image du point B par la projection canonique, sur l'espace $\overrightarrow{\mathcal{L}}$, de la droite vectorielle qui porte les points A et B, qui a pour origine le point A et qui est orientée de A vers B. Nous voyons

que cette distance est, non plus un nombre réel, mais un élément positif de l'espace vectoriel orienté $\overrightarrow{\mathcal{L}}$, de dimension 1.

Soit maintenant $\overrightarrow{\mathcal{L}}^2 = \overrightarrow{\mathcal{L}} \otimes \overrightarrow{\mathcal{L}}$ le produit tensoriel de deux exemplaires de $\overrightarrow{\mathcal{L}}$. C'est encore un espace vectoriel orienté de dimension 1, que nous appellerons *espace des produits euclidiens*. Soient $\overrightarrow{V}_1$ et $\overrightarrow{V}_2$ deux vecteurs non nuls, éléments de l'espace vectoriel $\overrightarrow{\mathcal{E}}$ associé à $\mathcal{E}$. Soit O un point quelconque de $\mathcal{E}$, A et B les points de $\mathcal{E}$ tels que $\overrightarrow{OA} = \overrightarrow{V}_1$ et $\overrightarrow{OB} = \overrightarrow{V}_2$. Il existe un déplacement euclidien unique g, élément du groupe G_E, qui laisse fixe le point O et qui applique la droite vectorielle qui porte les points O et A, qui a pour origine O et qui est orientée de O vers A, sur la droite vectorielle qui porte les points O et B, qui a pour origine O et qui est orientée de O vers B : c'est une rotation autour du point O, dont l'angle sera noté θ. Cet angle n'est déterminé qu'au signe près, et modulo 2π. Mais son cosinus est, quant à lui, parfaitement déterminé de manière unique. Nous pouvons donc définir le produit euclidien $g(\overrightarrow{V}_1, \overrightarrow{V}_2)$ des vecteurs $\overrightarrow{V}_1$ et $\overrightarrow{V}_2$ en posant

$$g(\overrightarrow{V}_1, \overrightarrow{V}_2) = \cos\theta |\overrightarrow{OA}| \otimes |\overrightarrow{OB}| \,.$$

Remarquons que $g(\overrightarrow{V}_1, \overrightarrow{V}_2)$ est, non un réel, mais un élément de l'espace des produits euclidiens $\overrightarrow{\mathcal{L}}^2$. En convenant que $g(\overrightarrow{V}_1, \overrightarrow{V}_2)$ est nul lorsque un au moins des deux vecteurs $\overrightarrow{V}_1$ ou $\overrightarrow{V}_2$ est nul, nous définissons une application $g : \overrightarrow{\mathcal{E}} \times \overrightarrow{\mathcal{E}} \to \overrightarrow{\mathcal{L}}^2$, appelée *produit euclidien*. Il est facile de voir que cette application est bilinéaire, symétrique et définie positive (ce qui signifie que pour tout vecteur non nul $\overrightarrow{V} \in \overrightarrow{\mathcal{E}}$, $g(\overrightarrow{V}, \overrightarrow{V})$ est un élément strictement positif de l'espace vectoriel orienté, de dimension 1, $\overrightarrow{\mathcal{L}}^2$).

Soit $\overrightarrow{l}$ un élément strictement positif de $\overrightarrow{\mathcal{L}}$. En prenant $\overrightarrow{l}$ comme vecteur unité de $\overrightarrow{\mathcal{L}}$ et $\overrightarrow{l} \otimes \overrightarrow{l}$ comme vecteur unité de $\overrightarrow{\mathcal{L}}^2$, nous pouvons identifier chacun de ces deux espaces à la droite réelle $\mathfrak{R}$. La distance $|\overrightarrow{AB}|$ de deux points A et B de $\mathcal{E}$, et le produit euclidien $g(\overrightarrow{V}_1, \overrightarrow{V}_2)$ de deux vecteurs éléments de $\overrightarrow{\mathcal{E}}$, peuvent alors être considérés comme des réels, et le produit euclidien peut être appelé *produit scalaire*, c'est-à-dire produit à valeurs dans le corps des scalaires $\mathfrak{R}$. Soucieux du sens physique qu'il faut atribuer aux objets mathématiques que nous manipulons, nous préférons cependant construire une théorie dans laquelle les lois physiques peuvent être formulées indépendamment de tout choix d'unités.

Appendice 2. Symboles de Christoffel et forces fictives

Considérons un point matériel soumis à un champ de forces f, mobile dans l'espace-temps non relativiste $\mathcal{U}$. Sa ligne d'univers est l'image d'une section $c : \mathcal{T} \to \mathcal{U}$ de la projection canonique $\tau : \mathcal{U} \to \mathcal{T}$ de l'espace-temps sur l'ensemble des temps $\mathcal{T}$, c'est-à-dire d'une application c qui vérifie $\tau \circ c = id_{\mathcal{T}}$. L'application c est solution de l'équation du mouvement sous forme intrinsèque

$$m \otimes \nabla_{Dc(t)}\big(Dc(t)\big) = f\big(c(t)\big) \,,$$

où $t \in \mathcal{T}$ est la variable indépendante (le temps), $Dc(t)$ la différentielle au point t de l'application c (c'est un élément de l'espace vectoriel $\mathcal{L}(\overrightarrow{\mathcal{T}}, \overrightarrow{\mathcal{U}})$ des applications linéaires de $\overrightarrow{\mathcal{T}}$ dans $\overrightarrow{\mathcal{U}}$), que l'on identifie canoniquement au produit tensoriel $\overrightarrow{\mathcal{T}}^* \otimes \overrightarrow{\mathcal{U}}$. Le terme $\nabla_{Dc(t)}\big(Dc(t)\big)$ est la dérivée covariante, pour la connexion affine de $\mathcal{U}$, de $Dc(t)$ relativement

à lui-même. C'est un élément de l'espace vectoriel $\overrightarrow{\mathcal{T}}^* \otimes \overrightarrow{\mathcal{T}}^* \otimes \overrightarrow{\mathcal{U}}$. Quant à m, c'est la masse du point matériel, élément de l'espace vectoriel de dimension 1 $\overrightarrow{\mu}$ des masses (nous ne le notons pas $\overrightarrow{\mathcal{M}}$ pour éviter toute confusion avec l'espace de Minkowski). La force f prend ses valeurs dans l'espace vectoriel $\overrightarrow{\mu} \otimes \overrightarrow{\mathcal{T}}^* \otimes \overrightarrow{\mathcal{T}}^* \otimes \overrightarrow{\mathcal{U}}$, et sa projection sur $\overrightarrow{\mu} \otimes \overrightarrow{\mathcal{T}}^* \otimes \overrightarrow{\mathcal{T}}^* \otimes \overrightarrow{\mathcal{T}}$, par l'application déduite de la projection canonique $\tau : \mathcal{U} \to \mathcal{T}$, est nulle (la composante temporelle de la force est toujours nulle).

Après le choix d'une unité de masse et d'une unité de temps, cette équation devient une équation différentielle, au sens usuel, dans l'espace affine $\mathcal{U}$, qui s'écrit

$$m \nabla_{\frac{dc(t)}{dt}} \frac{dc(t)}{dt} = f\big(c(t)\big)\,,$$

dont les deux membres peuvent être considérés comme prenant leurs valeurs dans l'espace vectoriel $\overrightarrow{\mathcal{U}}$.

Dans un système de coordonnées locales curvilignes quelconque (x^0, x^1, x^2, x^3) de $\mathcal{U}$, les composantes de la dérivée covariante $\nabla_{\frac{dc(t)}{dt}} \frac{dc(t)}{dt}$ ont pour expression (en écrivant $c^i(t)$ pour $x^i\big(c(t)\big)$, et en utilisant la convention d'Einstein de sommation par rapport aux indices répétés) :

$$\nabla_{\frac{dc(t)}{dt}} \frac{dc^i(t)}{dt} = \frac{d^2c^i(t)}{dt^2} + \Gamma^i_{j\,k} \frac{dc^j}{dt} \frac{dc^k}{dt}\,,$$

où les $\Gamma^i_{j\,k}$ sont les symboles de Christoffel de la connexion pour le système de coordonnées considéré. Les composantes de l'équation du mouvement s'écrivent donc :

$$m \left(\frac{d^2c^i(t)}{dt^2} + \Gamma^i_{j\,k} \frac{dc^j}{dt} \frac{dc^k}{dt} \right) = f^i\big(c(t)\big)\,, \quad i = 0\,,\ 1\,,\ 2\,,\ 3\,.$$

Supposons le système de coordonnées (x^0, x^1, x^2, x^3) tel que $x^0 = t$ soit le temps et que (x^1, x^2, x^3) soient les coordonnées d'espace dans un repère inertiel orthonormé. Les symboles de Christoffel associés à ce système sont tous identiquement nuls, et les équations ci-dessus deviennent :

$$m \frac{d^2c^i(t)}{dt^2} = f^i\big(c(t)\big)\,, \quad i = 0\,,\ 1\,,\ 2\,,\ 3\,,$$

(les f^i étant les composantes de la force f dans le repère considéré ; on a $f^0 = 0$ puisque la force doit avoir une projection nulle sur l'ensemble des temps).

Soit maintenant (y^0, y^1, y^2, y^3) un système de coordonnées dans $\mathcal{U}$ déduit du précédent par les formules

$$\begin{aligned} y^0 &= x^0 = t\,, & y^3 &= x^3\,, \\ y^1 &= -x^1 \cos(\omega x^0) + x^2 \sin(\omega x^0)\,, & y^2 &= -x^1 \sin(\omega x^0) + x^2 \cos(\omega x^0) \end{aligned}$$

(le second repère tourne, relativement au premier, autour de l'axe des x^3, à une vitesse angulaire constante ω). Soient $\overline{\Gamma}^\alpha_{\beta\,\gamma}$ les symboles de Christoffel de la connexion pour ce nouveau système. Les formules générales de transformation des symboles de Christoffel dans

un changement de système de coordonnées sont (toujours avec la convention d'Einstein de sommation par rapport aux indices répétés) :

$$\overline{\Gamma}^{\alpha}_{\beta\,\gamma} = \frac{\partial x^i}{\partial y^\beta \partial y^\gamma}\frac{\partial y^\alpha}{\partial x^i} + \Gamma^i_{j\,k}\frac{\partial x^j}{\partial y^\beta}\frac{\partial x^k}{\partial y^\gamma}\frac{\partial y^\alpha}{\partial x^i}\,.$$

On voit donc que les seuls symboles de Christoffel relativement au nouveau système de coordonnées éventuellement non nuls sont les $\overline{\Gamma}^{\alpha}_{\beta\,\gamma}$ pour $\alpha = 1$ ou 2, β et $\gamma = 0$, 1 ou 2. On obtient aisément

$$\begin{array}{lll} \overline{\Gamma}^1_{0\,0} = -\omega^2 y^1\,, & \overline{\Gamma}^1_{1\,0} = \overline{\Gamma}^1_{0\,1} = 0\,, & \overline{\Gamma}^1_{2\,0} = \overline{\Gamma}^1_{0\,2} = -\omega\,, \\ \overline{\Gamma}^2_{0\,0} = -\omega^2 y^2\,, & \overline{\Gamma}^2_{1\,0} = \overline{\Gamma}^2_{0\,1} = \omega\,, & \overline{\Gamma}^2_{2\,0} = \overline{\Gamma}^2_{0\,2} = 0\,. \end{array}$$

Les équations du mouvement du point matériel dans les coordonnées (y^0, y^1, y^2, y^3) s'écrivent donc :

$$\begin{array}{ll} m\dfrac{d^2y^0}{dt^2} & = 0\,, \\ m\dfrac{d^2y^1}{dt^2} & = m\omega^2 y^1 + 2m\omega\frac{dy^2}{dt} + \overline{f}^1\big(c(t)\big)\,, \\ m\dfrac{d^2y^2}{dt^2} & = m\omega^2 y^2 - 2m\omega\frac{dy^1}{dt} + \overline{f}^2\big(c(t)\big)\,, \\ m\dfrac{d^2y^3}{dt^2} & = \overline{f}^3\big(c(t)\big)\,, \end{array}$$

où les $\overline{f}^i$ ($i = 1$, 2 ou 3) sont les composantes de la force f dans le nouveau repère.

Au second membre de la seconde et de la troisième équations de ce système, on voit apparaître les composantes de la force centrifuge $(m\omega^2 y^1, m\omega^2 y^2)$ et de la force de Coriolis $(2m\omega\frac{\partial y^2}{\partial t}, -2m\omega\frac{\partial y^1}{\partial t})$. Ces forces, considérées en Mécanique comme des "forces fictives", sont tout simplement les termes contenant les symboles de Christoffel dans l'expression de la dérivée covariante.

Plus généralement, E. Cartan a montré [[22] a] qu'en modifiant la connexion de $\mathcal{U}$ on pouvait transformer un champ de forces que les mécaniciens considèrent en général comme "réel", comme par exemple le champ de pesanteur, afin de l'inclure dans la géométrie de l'espace-temps. Ce champ de forces n'apparaîtra alors, dans les équations du mouvement explicitées en coordonnées locales, que par les symboles de Christoffel.

Appendice 3. Structure intrinsèque de l'espace-temps de Minkowski

Nous avons vu que l'espace de Minkowski $\mathcal{M}$ est un espace affine réel de dimension 4, sur lequel agit le groupe de Poincaré G_P. Soit O un point quelconque de $\mathcal{M}$. En utilisant ce point comme origine nous identifions $\mathcal{M}$ avec son espace vectoriel associé $\overrightarrow{\mathcal{M}}$. Le sous-groupe de G_P formé par les éléments qui laissent le point O fixe s'identifie au groupe de Lorentz G_L. On sait qu'il existe sur $\overrightarrow{\mathcal{M}}$ des formes quadratiques qui sont invariantes par l'action du groupe de Lorentz. Ces formes, mises sous forme réduite, s'expriment comme des sommes algébriques de quatre carrés, dont un est affecté d'un signe et les trois autres

du signe opposé ; elles sont toutes proportionnelles à l'une quelconque d'entre elles, pourvu que cette dernière soit non nulle. En d'autres termes, elles forment un espace vetoriel de dimension 1, que nous noterons $\overrightarrow{\mathcal{Q}}$. Soit $\overrightarrow{\mathcal{Q}}^*$ son dual. Nous définissons une application bilinéaire symétrique $g : \overrightarrow{\mathcal{M}} \times \overrightarrow{\mathcal{M}} \to \overrightarrow{\mathcal{Q}}^*$ en posant, pour toute forme quadratique $\eta \in \overrightarrow{\mathcal{Q}}$ et tout couple $(\overrightarrow{V}_1, \overrightarrow{V}_2)$ d'éléments de $\overrightarrow{\mathcal{M}}$:

$$\langle g(\overrightarrow{V}_1, \overrightarrow{V}_2), \eta \rangle = \eta(\overrightarrow{V}_1, \overrightarrow{V}_2)\,,$$

où le crochet $\langle\ ,\ \rangle$ au membre de gauche est le couplage par dualité de $\overrightarrow{\mathcal{Q}}^*$ avec $\overrightarrow{\mathcal{Q}}$ et où, au membre de droite, η désigne la forme bilinéaire symétrique associée à la forme quadratique, élément de $\overrightarrow{\mathcal{Q}}$, elle aussi notée η. L'application g ainsi définie sera appelée *structure de Minkowski intrinsèque* sur l'espace $\mathcal{M}$.

Remarques

1 La construction décrite ci-dessus pourrait être appliquée à l'espace affine réel $\mathcal{E}$, de dimension 3, utilisé comme modèle de l'espace physique. Cela donnerait une autre construction de l'espace $\overrightarrow{\mathcal{L}}^2$ des produits euclidiens, équivalente à celle donnée dans l'Appendice 1.

2 Le choix d'un élément non nul η de $\overrightarrow{\mathcal{Q}}$ détermine, dans l'espace $\mathcal{M}$, à la fois une unité de longueur et une unité de temps. Supposons par exemple que la forme quadratique η, mise sous forme canonique, comporte un carré affecté du signe + et trois carrés affectés du signe − (nous dirons qu'elle est de signature $(+,-,-,-)$). Un vecteur $\overrightarrow{V} \in \overrightarrow{\mathcal{M}}$ sera dit *de genre temps* si $\eta(\overrightarrow{V}, \overrightarrow{V}) > 0$, *de genre espace* si $\eta(\overrightarrow{V}, \overrightarrow{V}) < 0$ et *isotrope* si $\eta(\overrightarrow{V}, \overrightarrow{V}) = 0$.

Si $\overrightarrow{V}$ est de genre temps, l'intervalle de temps qu'il représente est mesuré, avec l'unité de temps associée à η, par le réel positif $\sqrt{g(\overrightarrow{V}, \overrightarrow{V})}$. S'il est de genre espace, la longueur qu'il représente est mesurée, avec l'unité de longueur associée à η, par le réel positif $\sqrt{-g(\overrightarrow{V}, \overrightarrow{V})}$.

3 Le groupe de Poincaré (même si on le restreint à sa composante neutre) agit transitivement sur chacun des trois ensembles de droites vectorielles orientées plongées dans l'espace de Minkowski $\mathcal{M}$:

– – l'ensemble des droites orientées de genre espace,
– – l'ensemble des droites de genre temps orientées vers le futur,
– – l'ensemble des droites isotropes orientées vers le futur.

Un élément du groupe de Poincaré qui laisse globalement invariante une droite vectorielle orientée plongée dans $\mathcal{M}$, restreint à cette droite, est l'application identique si cette droite est de genre espace, ou de genre temps, mais pas si cette droite est isotrope. La méthode utilisée dans l'Appendice 1 permet donc de comparer entre eux des segments de deux droites différentes, pourvu que ces droites soient toutes deux de genre espace, ou toutes deux de genre temps, et de définir, pour les droites de genre espace d'une part, et pour les droites de genre temps d'autre part, un espace vectoriel orienté de dimension 1 dans lequel la mesure de ces intervalles prend ses valeurs.

Par contre, on ne peut en général pas comparer des intervalles de deux droites isotropes différentes.

Dans la plupart des ouvrages d'enseignement traitant de la théorie de la Relativité, il est question de "contraction des longueurs" et de "dilatation des temps". Or ce qui précède montre bien qu'on peut comparer entre eux des segments de droite appartenant à deux droites différentes de genre espace, ou à deux droites différentes de genre temps. Lorsqu'on raisonne dans l'espace-temps, sur les événements, il n'y a pas de contraction des longueurs, ni de dilatation des temps. Ces concepts, imaginés par FitzGerald et Lorentz pour expliquer le résultat négatif des expériences de Michelson et Morley avant qu'Einstein ne crée la théorie de la Relativité, sont à l'origine de bien des confusions et devraient être, sinon bannis des ouvrages d'enseignement, du moins présentés différemment, en suivant le développement historique de la théorie.

4 Les opérations mathématiques décrites dans l'Appendice 1 correspondent exactement aux opérations physiques que l'on fait pour mesurer une distance, en plaçant une règle graduée le long de l'objet à mesurer. Les opérations mathématiques décrites dans le présent appendice correspondent aux opérations physiques que l'on fait pour mesurer une distance en mesurant le temps mis par la lumière pour la parcourir. Historiquement, ces deux manières de procéder ont été utilisées successivement pour définir le mètre : d'abord défini comme la quarante millionième partie du méridien terrestre, puis (de 1889 à 1960) comme la distance entre les deux traits gravés sur le mètre étalon déposé au pavillon de Breteuil, puis par référence à la longueur d'onde d'une raie spectrale du krypton 86, le mètre est, depuis 1983, la distance parcourue par la lumière dans le vide pendant une durée de 1/299792458 secondes.

Bibliographie

[1] V.I. Arnold, *Mathematical methods of classical mechanics*, second edition, Springer-Verlag, New York, 1989 ; première édition en russe, Nauk, Moscou, 1974.

[2] H. Bergson, *Durée et simultanéité, à propos de la théorie d'Einstein*, Librairie Félix Alcan, Paris, 1922 ; réédité par Quadrige / Presses Universitaires de France, Paris, 1968.

[3] H. Cabannes, *Cours de mécanique générale*, Dunod, Paris, 1962.

[4] É. Cartan, (a) *Les variétés à connexion affine et la théorie de la relativité généralisée*, I et II, Ann. Ec. Norm. **40**, 1923, pp. 325–342 et **41**, 1924, pp. 1–25 (se trouve aussi dans les *Œuvres complètes*, partie III 1, pp. 659–746 et 799-823). (b) *Œuvres complètes*, parties III 1 et III 2, pp. 747–797, 825–862, 863–889, 1239–1243, Éditions du CNRS, Paris, 1984.

[5] É. Cartan, A. Einstein, *Letters on absolute parallelism, 1929–1932*, Princeton University Press et Académie Royale de Belgique, Princeton, 1979.

[6] R. Dugas, *Histoire de la Mécanique*, Éditions du Griffon, Neuchâtel, 1950. Réimprimé par les Éditions Jacques Gabay, Paris, 1996.

[7] Ch. Ehresmann, *Les connexions infinitésimales* , Colloq. Topologie (Bruxelles 1950). Œuvres complètes, Partie I 2, pp. 179–205.

[8] A. Einstein, (a) *Sur l'électrodynamique des corps en mouvement*, traduit par M. Solovine d'après l'article des Annalen der Physik, XVII, 1905, Gauthier-Villars, Paris, 1925. (b) *L'éther et la théorie de la Relativité* et *La Géométrie et l'expérience*, Gauthier-Villars, Paris, 1953. (c) *Quatre conférences sur la théorie de la Relativité faites à l'université de Princeton* traduites par M. Solovine, Gauthier-Villars, Paris, 1925 (également publiées en anglais en un volume sous le titre *The Meaning of Relativity*, Princeton University Press, Princeton, 1922). (d) *Signification de la Relativité, compléments*, traduit par M. Solovine et M.-A. Tonnelat, Gauthier-Villars, 1960. (e) *Théorie de la gravitation généralisée*, traduit par M. Solovine d'après le second appendice de la troisième édition de *The Meaning of Relativity*, 1950. Textes rassemblés en un seul volume et réimprimés par les Éditions Jacques Gabay, Paris, 1994.

[9] M. Friedman, *Foundations of Space-Time Theories*, Princeton University Press, Princeton, 1983.

[10] Hergé, *L'affaire Tournesol*, page 45, et *Vol 714 pour Sydney*, page 33. Casterman, Paris-Tournai, 1956 et 1968.

[11] J.-L. Lagrange, *Méchanique analitique*, chez la veuve Desaint, Paris, 1788. Réimprimé par les Éditions Jacques Gabay, Paris, 1989.

[12] P. Libermann, *Cartan connections and momentum maps*, à paraître dans les Banach Center Publications, Warszawa.

[13] A. Lichnerowicz, (a) *Variétés symplectiques, variétés canoniques et systèmes dynamiques*, dans *Topics in Differentiable geometry*, pp. 57–85, Academic Press, New York, 1976. (b) *Variétés symplectiques et variétés canoniques*, dans *Trends in applications of pure mathematics to mechanics*, pp. 249–261, Pitman, London, 1976. (c) *New geometrical dynamics*, in *Differential geometrical methods in Mathematical Physics*, pp. 377–394, Lecture notes in Mathematics **570**, Springer Verlag, Berlin, 1977. (d) *Les variétés de Poisson et leurs algèbres de Lie associées*, J. Differential Geometry **12** (1977), pp. 253–300. (e) *Les variétés de Jacobi et leurs algèbres de Lie associées*, J. Math. Pures et Appl. **57** (1978), pp. 453–488. (f) *La géométrie des transformations canoniques*, Bull. Soc. math. de Belgique **31** (1979), pp. 105–135.

[14] H. A. Lorentz, (a) *Versuch einer Theorie der elektrischen und optischen Erscheinungen in bewegten Körpen*, Leyde, 1895. (b) *Electromagnetic phenomena in a system moving with any velocity smaller than that of light*, Proc. Acad. Sc. Amsterdam, **6** (1904), p. 809.

[15] E. Mach, *La Mécanique, exposé historique et critique de son développement*, ouvrage traduit de la quatrième édition allemande par E. Bertrand, Hermann, Paris, 1904. Réimprimé par les Éditions Jacques Gabay, Paris, 1987.

[16] J. Mandel, *Cours de mécanique, tome I*, École Polytechnique, Année scolaire 1953–1954, Paris.

[17] A. A. Michelson, E. W. Morley, American Journal of Science, **22** (1887), p. 333.

[18] R. L.Minkowski, Conférence faite à Cologne le 21 septembre 1908, citée et analysée dans le livre de René Dugas [6], pp. 468–473.

[19] I. Newton, *Principes mathématiques de la Philosophie naturelle*, tomes I et II, traduit par Madame la Marquise du Chastellet, chez Desaint et Saillant, Paris, 1759. Réimprimé par les Éditions Jacques Gabay, Paris, 1990.

[20] F. Nietzsche, *Le gai savoir* et *Ainsi parlait Zarathoustra*, *Œuvres philosophiques complètes*, volumes V et VI, Gallimard, Paris, 1971 et 1982.

[21] P. Painlevé*Les axiomes de la Mécanique, examen critique, avec une note sur la propagation de la lumière*, Gauthier-Villars, Paris, 1922. Réimprimé par les Éditions Jacques Gabay, Paris, 1995.

[22] J. Pérès, *Mécanique générale*, deuxième édition, Masson, Paris (1962).

[23] H. Poincaré, (a) *La science et l'hypothèse*, Ernest Flammarion, Paris, 1906. (b) *Science et méthode*, Ernest Flammarion, Paris, 1908. (c) *La valeur de la science*, Ernest Flammarion, Paris, 1912. (d) *Dernières pensées*, Ernest Flammarion, Paris, 1913. (e) *La Mécanique nouvelle*, livre réunissant en un seul volume le texte d'une conférence faite au congrès de Lille de l'Association française pour l'avancement des sciences en 1909, le mémoire du 23 juillet 1905 intitulé *Sur la dynamique de l'électron*, publié aux Rendiconti del Circolo matematico di Palermo **XXI** (1906) et une Note aux Comptes Rendus

de l'Académie des Sciences, de même titre (séance du 15 juin 1905, **CXL**, 1905, p. 1504) ; Gauthier-Villars, Paris, 1924 ; réimprimé par les Éditions Jacques Gabay, Paris, 1989. (f) *Les méthodes nouvelles de la mécanique céleste*, tomes I, II et III, Gauthier-Villars, Paris, 1892, 1893, 1899, réimprimé par la Librairie scientifique et technique Albert Blanchard, Paris, 1987.

[24] A. Simaan, J. Fontaine, *L'image du monde des Babyloniens à Newton*, AdaptÉditions, Paris, 1998.

[25] J.-M. Souriau, *Structure des systèmes dynamiques*, Dunod, Paris, 1970.

[26] W. M. Tulczyjew, *Geometric Formulations of Physical Theories*, Bibliopolis, Napoli, 1989.

[27] B. C. van Frassen, *An introduction to the Philosophy of Space and Time*, Columbia University Press, 1970.

[28] W. Voigt, Göttinger Nachrichten (1887), p. 41. Cité dans le livre de R. Dugas [6], page 454.

[29] A. Weinstein, (a) *The local structure of Poisson manifolds*, J. Diff. Geom. **18** (1983), pp. 523–557. (b) *Symplectic groupoids and Poisson manifolds*, Bull. Amer. Math. Soc. **16** (1987), pp. 101–104.

[30] H. Weyl, *Temps, espace, matière*, traduit de la quatrième édition allemande par G. Juvet et R. Leroy, 1922 ; réimprimé par la Librairie scientifique Albert Blanchard, Paris, 1958.

The Geometry of Momentum

Alan Weinstein[1]
Department of Mathematics
University of California
Berkeley, CA 94720 USA
(alanw@math.berkeley.edu)

To Jerry Marsden for his 60th birthday

1 Introduction

The term *momentum* (in Italian, *momento*) was introduced by Galileo Galilei as the "virtue" of a moving object which keeps it moving. He expressed it as the product of weight and velocity. Galileo's notion of momentum was an outgrowth of the medieval *impetus* theory of William of Occam, Jean Buridan, Nicole Oresme, and others (see, for example, Dugas [9]), which deviated from Aristotle's view that something external to a moving (terrestrial) body was necessary to keep it in motion. (Opposition to Aristotle's view had also been expressed as early as the 5th century by John (Philopon) of Alexandria.)

By now, "momentum" has a much wider meaning. Many applications of the term refer to quantities whose conservation under the time evolution of a physical system is related to some symmetry of the system. The theme of this paper is this relation between momentum and symmetry, especially as the meaning of the later was extended in the late 20th century. Though much of this extension was motivated by quantum mechanics, the emphasis here will be on momentum in classical mechanics. Another very important aspect of our subject which we will not treat in this paper is *symplectic reduction*, which is the simultaneous use of symmetries and conserved quantities to reduce the dimensionality of a hamiltonian system. We refer to [17] for an extensive history of reduction.

[1]Research partially supported by NSF Grant DMS-99-71505.
Keywords : moment map, momentum map,hamiltonian, symplectic, symmetry, conservation law, groupoid.
MSC2000 Subject Classification Numbers : 53D20 (Primary), 20L05, 53D17, 70H33 (Secondary).

2 Momentum in the calculus of variations

Despite the original conception of momentum as *proportional* to velocity, we now see momentum for the most part as *dual* to velocity. This idea appears clearly in the calculus of variations where, given a lagrangian $L(t, x, \dot{x})$ for a function $x = (x^1, ..., x^n)$ of t, we define the conjugate momenta by $p_i = \partial L/\partial \dot{x}^i$ and then the hamiltonian by $H = \sum p_i \dot{x}^i - L$. One sees from the way in which the conjugate momenta transform under coordinate changes (which results in the invariance of H under such changes) that the conjugate momenta are the components of a vector p which is dual to the velocity vector $\dot{x}$.

Of course, if $L = \frac{1}{2} m \sum_i (\dot{x}^i)^2 - V(x)$, then $p_i = m\dot{x}^i$. The conflict between lower and upper indices here results from the use of orthonormal coordinates. More invariantly, we should write the first (kinetic energy) part of L as

$$\frac{1}{2} \sum_{i,j} m_{ij} \dot{x}^i \dot{x}^j, \tag{1.6}$$

where $m_{ij} = m\delta_{ij}$. The "mass tensor" m_{ij} may then replaced in (1.6) by an arbitrary riemannian metric $m_{ij}(x)$ to describe in curvilinear coordinates what Smale [21] calls a *simple mechanical system*. For such a system, the relation between momentum and velocity becomes

$$p_i = \sum_j m_{ij} \dot{x}^j, \tag{1.7}$$

the mass tensor now being interpreted as a mapping from velocity (tangent) to momentum (cotangent) vectors.

If we add the vector potential of a magnetic field to obtain

$$L = \frac{1}{2} \sum_{i,j} m_{ij} \dot{x}^i \dot{x}^j + e \sum_i A_i(x) \dot{x}^i + V(x),$$

then the momentum and velocity are no longer proportional; instead,

$$p_i = \sum_j m_{ij} \dot{x}^j + A_i(x). \tag{1.8}$$

To see why (1.8) and not the simpler relation (1.7) is the appropriate definition of momentum here, let us consider the example in 3-dimensional configuration space where $m_{ij} = m\delta_{ij}$, the vector potential is $ex^1\dot{x}^3$, and $V = 0$. This lagrangian is invariant under translations of configuration space in the direction of the x^3-axis; if we look for conserved quantities, we find that $p_3 = m\dot{x}^3 + ex^1$ is a constant of motion, while the single term $m\dot{x}^3$ is not. (The motions are circular helices with axis in the x^2 direction.)

Investigating this example further, one finds that $p_2 = m\dot{x}^2$ is also a conserved quantity, while $p_1 = m\dot{x}^1$ is not. In fact, translations along the x^1 axis are not symmetries of the lagrangian; however, one may check directly from the equations of motion that there is a third conserved quantity, namely $m\dot{x}^1 - ex^3$, found most easily from the hamiltonian formalism, as follows. The hamiltonian $H = L - \sum_i p_i \dot{x}^i$ may be written in terms of positions and velocities as $\frac{m}{2} \sum_i (\dot{x}^i)^2$ or in terms of positions and conjugate momenta as $\frac{1}{2m}(p_1^2 + p_2^2 +$

$(\frac{p_3}{m} - \frac{e}{m}x^1)^2)$. From this formula, one sees that $\frac{\partial}{\partial x^1} + e\frac{\partial}{\partial p_3}$ is a symmetry; its generating hamiltonian is $-ex^3 + p^1 = m\dot{x}^1 - ex^3$.
Symplectic geometry gives a way to recover the proportionality of momentum and velocity, with the additional gain of "gauge invariance," (i.e., the magnetic field and not the potential enters the equations). One introduces the coordinates $\xi_i = m\dot{x}^i$, so that the hamiltonian has the simple form $\frac{1}{2m}\sum_i(\xi_i)^2$, but the symplectic form $\sum dx^i \wedge dp_i$ becomes in these coordinates $\sum_i dx^i \wedge d\xi_i + e\,dx^3 \wedge dx^1$. The x^i and ξ_i are no longer canonically conjugate variables; the extra term $e\,dx^3 \wedge dx^1$ in the symplectic form is the magnetic field. In this formulation, the symmetry of the system with respect to rotation around the x^2 axis becomes manifest.

In the language of the late 20th century, one says that the phase space of a particle in a magnetic field is no longer the cotangent bundle of configuration space; rather it is a "shifted cotangent bundle" obtained by beginning with the cotangent bundle of a principal circle bundle over configuration space and then applying the operation of symplectic reduction. (See [3] for an exposition.)

Another link between this example and 20th century mathematics comes through the work of Noether [19], who, in the course of developing ideas of Einstein and Klein in general relativity theory, found a very general equivalence between symmetries and conservation laws in field theory (i.e. variational problems which may involve several independent variables) now known as *Noether's theorem*. Noether's work on this subject is so important that her name has become inextricably attached to the relation between symmetry and conservation laws. It would be interesting to trace the early history of this relation, perhaps even as far as ancient Greek astronomy, where the uniform circular motion of heavenly bodies might have been related to the perfection (i.e. rotational symmetry) of the celestial sphere.

In the remainder of this paper, we will concentrate on the hamiltonian point of view toward "Noether's theorem."

3 Hamiltonian systems with symmetry

The relation between conserved quantities and symmetries becomes very simple when expressed in terms of Poisson brackets. On a symplectic manifold P, the time evolution of a function F under the hamiltonian flow of a function H is given by the formula

$$\frac{dF}{dt} = \{F, H\}.$$

Antisymmetry of this bracket implies immediately that F is a conserved quantity for the hamiltonian flow of H if and only if the hamiltonian flow of F consists of symmetries of H. If F and G are both conserved quantities, then so is $\{F, G\}$, by Poisson's theorem, which is most conveniently proved via the Jacobi identity. If $F_1, \ldots, F_k$ are conserved quantities, then so is any function $g(F_1, \ldots, F_k)$, so the process of building new conserved quantities from old ones naturally terminates if we arrive at a list $F_1, \ldots, F_k$ of functions such that, for each pair of functions in the list, $\{F_i, F_j\} = \pi_{ij}(F_1, \ldots F_k)$ for functions π_{ij}. Lie [13] refers to such a list of functions as generating a *function group*, the function group itself consisting of all the $g(F_1, \ldots, F_k)$. If $F_1, \ldots, F_k$ are functionally independent, then the π_{ij} are uniquely

determined smooth functions on $\mathbb{R}^k$ and define on the set of all smooth functions on $\mathbb{R}^k$ the structure of what Lie [13] calls an *abstract function group*. This is a Lie algebra structure defined by the bracket

$$\{F, G\} = \sum_{i,j} \pi_{ij}(\mu_1, \ldots, \mu_k) \frac{\partial F}{\partial \mu_i} \frac{\partial G}{\partial \mu_j}.$$

An important special case is that where the functions π_{ij} are linear. The linear functions on $\mathbb{R}^k$ are then closed under the bracket operation and form a Lie algebra $\mathfrak{g}$; the carrier $\mathbb{R}^k$ of the abstract function group is identified with the dual space $\mathfrak{g}^*$.

In geometric terms, a phase space with a symmetry group consists of a manifold P equipped with a symplectic structure ω and a hamiltonian action of a Lie group G. By the latter, we mean a symplectic action of G on P together with an equivariant map J from P to the dual $\mathfrak{g}^*$ of the Lie algebra of G such that, *for each $v \in \mathfrak{g}$, the 1-parameter group of transformations of P generated by v is the flow of the hamiltonian vector field with hamiltonian $x \mapsto \langle J(x), v \rangle$*. The map J is called the *momentum map* (or, by many authors, *moment map*) of the hamiltonian action. If one is simply given a symplectic action of G on P, any map J satisfying the condition in italics above, even if it is not equivariant, is called *a* momentum map for the action.

The antisymmetry of the Poisson bracket now implies the hamiltonian version of Noether's theorem : *the momentum map J is invariant under the hamiltonian flow of H if and only if H is invariant under the action of the component of the identity in the group G.* We defer until later (see Sections 6 and 7) the question of how nonidentity components of a symmetry group G are related to conservation laws.

The duality between symmetries and conservation laws in the hamiltonian formalism becomes most striking in the language of *Poisson geometry*, which is the contemporary name for the geometry of Lie's abstract function groups. We recall that a *Poisson manifold* is a manifold M with a Lie algebra structure $\{\ ,\ \}$ on its algebra of smooth functions which is an algebra derivation in each argument. For a function H, the derivation $F \mapsto \{F, H\}$ is the *hamiltonian vector field* associated to H. If a group G acts by automorphisms of a symplectic manifold P in such a way that the quotient space P/G inherits a manifold structure, then P/G becomes a Poisson manifold through the identification of the smooth functions on P/G with the G-invariant smooth functions on P.[2] The natural projection $\pi : P \to P/G$ is a *Poisson map* in the sense that its pullback on functions is a Lie algebra homomorphism. If the action of G is hamiltonian, its equivariant momentum map $J : P \to \mathfrak{g}^*$ is also a Poisson map, and the diagram of Poisson maps $\mathfrak{g}^* \leftarrow P \to P/G$ is called a *dual pair*. In symplectic terms, this means that the tangent spaces of the fibres of the two maps are symplectically orthogonal to one another. The algebras of functions on $\mathfrak{g}^*$ and P/G commute with one another as subalgebras of the Lie algebra of functions on P, and if the fibres of the two maps in the dual pair are connected, these two subalgebras are the full commutants of one another.

As we have suggested above, many of the ideas in this section can already be found in Lie's book [13]. (See [24] for a brief historical discussion.)

[2] If P/G is not a manifold, we may treat it as some kind of "Poisson variety" or "Poisson space."

4 Symplectic groupoid actions

The "duality" between $\mathfrak{g}^*$ and P/G would be more symmetrical if the fibres of J could also be seen as orbits of an action. In general, P/G is not the dual of a Lie algebra, but there is an object which plays the role of the second group in this picture; it is a *symplectic groupoid* for the Poisson manifold P/G. This is a groupoid[3] whose objects are the points of P/G and whose morphisms form a symplectic manifold $\Gamma(P/G)$. The symplectic structure on $\Gamma(P/G)$ is compatible with the Poisson structure on P/G in the sense that target and source maps α and β from $\Gamma(P/G)$ to P/G are Poisson and antiPoisson (i.e. reversing the sign of Poisson brackets) respectively. The groupoid multiplication is compatible with the symplectic structure in the following sense. Writing Γ for $\Gamma(P/G)$ and denoting by $\Gamma^{(2)} \subset \Gamma \times \Gamma$ the manifold of composable pairs, we have three natural maps from $\Gamma^{(2)}$ to Γ, namely the cartesian product projections p_1 and p_2 and the groupoid multiplication m. The symplectic structure ω on Γ should then satisfy the cocycle condition $p_1^*(\omega) - m^*(\omega) + p_2^*(\omega) = 0$.

The quotient projection $\pi : P \to P/G$ may be seen as a momentum map (in an extended sense) for a symplectic groupoid action of $\Gamma(P/G)$ on P whose orbits are the fibres of J. This means (see [18] that we have a map $a : \Gamma \times_{P/G} P \to P$ satisfying the usual "associativity property" of an action; the action being symplectic means that $p_1^*(\omega) - a^*(\omega_P) + p_2^*(\omega_P) = 0$. The fibre product here is taken with respect to the maps β and π, ω_P is the symplectic structure on P, and p_1 and p_2 are again the cartesian projections from $\Gamma \times_{P/G} P$ to Γ and P. If we write gx for $a(g, x)$, then we have the equivariance relation $J(gx) = gJ(x)$ involving the natural action of the groupoid on its objects. We call π the *momentum map* of the groupoid action.

When $\mathfrak{g}$ is the Lie algebra of G, a natural choice for $\Gamma(\mathfrak{g}^*)$ is the cotangent bundle T^*G with its canonical symplectic structure and a groupoid structure whose source and target maps to $\mathfrak{g}^* = T_e^*G$ are given by left and right translation. Under the identification of T^*G with $\mathfrak{g}^* \times G$ by left translations, the groupoid multiplication is given by $(\mu, g)(\nu, h) = \pi(\nu, gh)$. The fact that the fibres of $\pi : P \to P/G$ are the orbits of the G action follows from the fact that the cotangent bundle projection is a groupoid homomorphism from $\Gamma(\mathfrak{g}^*) = T^*G$ to G (considered as a groupoid with a single object), and that the symplectic groupoid action on P factors through this homomorphism. For a more general Poisson manifold like P/G, there is no such homomorphism from $\Gamma(P/G)$ to a group, and so the symplectic groupoid is absolutely essential to the picture.

Passing from groups to groupoids leads to a new definition of a dual pair. We may define a dual pair to consist of a pair of symplectic groupoids Γ_1 and Γ_2 with underlying Poisson manifolds Q_1 and Q_2, together with commuting symplectic groupoid actions of Γ_1 and Γ_2 on a symplectic manifold P having momentum maps J_1 and J_2 such that the orbits of each action are the fibres of the momentum map of the other. We note that this formulation includes the case where all the groupoids and spaces involved are discrete. The groupoid action contains both aspects of the symmetry/conservation duality, since the conserved quantity is just the momentum map of the groupoid action.

Our main point, then, is that a hamiltonian action of a Lie group should be seen as

[3]We recall that a *groupoid* is a small category whose morphisms are all invertible; if this category has just one object, the groupoid is a group.

a special case of a more flexible notion of symmetry–the action of a symplectic groupoid. Such an action always comes equipped with a momentum map which is a Poisson map and which in fact determines the action of the component of the groupoid containing the identity elements. Pursuing this idea, it is natural to ask whether properties of momentum maps of hamiltonian group actions extend to the groupoid case. As a first step toward answering this question, we have established in [26] an extension of Kirwan's nonabelian version of Atiyah, Guillemin, and Sternberg's convexity theorem.

5 Poisson Lie groups and beyond

We have seen that the cotangent bundle of a Lie group G is a rather special symplectic groupoid in that it admits a homomorphism to a group. This homomorphism, the cotangent bundle projection π, is a covering morphism in the sense of [4]; this means that its fibres are *bisections* of the groupoid, i.e. they project diffeomorphically to the space of objects under the source and target maps. Finally, π is a Poisson map when G carries the zero Poisson structure. This is equivalent to the fact that the fibres of π are lagrangian submanifolds of the cotangent bundle.

The properties of the symplectic groupoids T^*G listed in the paragraph above nearly characterize these examples, the only exceptions being the quotients of cotangent bundles by lattices of bi-invariant (and hence closed) 1-forms. On the other hand, if we drop the requirement that the fibres of a covering morphism π from a symplectic groupoid Γ to a group G be lagrangian, we obtain a much larger class of symplectic groupoids which are still closely connected to groups. The group G must now carry a nonzero Poisson structure in order for π to be a Poisson map; compatibility of the group structure with the Poisson structure is expressed by the condition that the multiplication $G \times G \to G$ be a Poisson map. A symplectic groupoid action of Γ on P still corresponds to an action of the group G on P. The manifold of objects of Γ, which is the target of the momentum maps of such actions, turns out to be a group itself; it is called the dual of the Poisson-Lie group G and is denoted G^*. (The group structure on G^* is essentially encoded in the Poisson structure on G, and vice-versa.) The momentum map is equivariant with respect to an action of G on G^* known as the *dressing action*. Since the fibres of π are not lagrangian, the elements of G no longer act on P by symplectic transformations. Rather, the action as a whole is a *Poisson action* in the sense that the map $G \times P \to P$ is a Poisson map. (The dressing action is also a Poisson action.) The original groupoid Γ is simultaneously a symplectic groupoid for both G and G^*; as a symplectic manifold, it is known as the *Heisenberg double* of the pair (G, G^*) of dual Poisson Lie groups. What we have described here is essentially Lu's momentum map theory [16] for Poisson Lie group actions.

Having dealt with compatibility conditions for symplectic structures on groupoids and Poisson structures on groups, one might ask for a common framework for both. This is provided by the notion of *Poisson groupoid* [25]. Since Poisson structures correspond to possibly degenerate bivector fields, one might hope for a theory which also includes degenerate 2-forms. This is provided by the theory of Dirac structures [8]. These are subbundles of a direct sum $TP \oplus T^*P$ which are maximal isotropic for a natural symmetric bilinear form and which are closed under a bracket discovered by Courant [8] and which has become the

prototype for an object known as a *Courant algebroid* [14].

Computations on moduli spaces of flat connections in gauge theory have led to a yet more general notion of hamiltonian symmetry. Here, the 2 form ω_P on phase space is neither closed nor nondegenerate, but these "defects" are compensated for by the presence of an auxiliary structure on the group, just as the noninvariance of the symplectic structure on a Poisson G-space is compensated for by the Poisson structure on the group. In the theory of *quasi-hamiltonian G-spaces*, introduced in [1], the symmetry group G acting on P now carries a bi-invariant inner product and, hence, a bi-invariant 3-form ϕ defined by $\phi(u,v,w) = \langle [u,v], w\rangle$. The momentum map is now *group-valued*, i.e. we have $J : P \to G$ with the property that $d\omega_P = J^*\phi$. As in the ordinary hamiltonian case, ω_P is G-invariant, and J is now equivariant with respect to the adjoint action of G on itself. The relation between the momentum map and the infinitesimal generators of the symmetry group is now more complicated. For $v \in \mathfrak{g}$, the corresponding vector field v_P on P satisfies the condition that,

$$i_{v_P}\omega_P = J^*\langle \frac{1}{2}(\theta + \overline{\theta}), v\rangle,$$

where θ and $\overline{\theta}$ are the left-invariant and right-invariant Maurer-Cartan forms on G. Note that, since ω_P may be degenerate, this condition does not necessarily determine the action uniquely for a given momentum map.

The relation between momentum and symmetry becomes even looser in Karshon's theory of "abstract momentum maps" [12]. Here, G is a torus acting on P, and J is an equivariant map from P to $\mathfrak{g}^*$. There is no longer a 2-form on P. Instead, the action and the momentum map are related by the condition that, along the fixed point set of any subgroup $H \subseteq G$, the image of J must lie in a subspace parallel to $\mathfrak{h}^\perp \subseteq \mathfrak{g}^*$; i.e. the component of the momentum map given by each element of $\mathfrak{h}$ is a conserved quantity for the restricted action. Remarkably, this weak relation is sufficient to prove versions of many results usually associated with ordinary momentum maps.

6 Universal momentum maps

There have been at least two quite different approaches to the idea of a "universal momentum map." Evans and Lu [10] construct for each Poisson Lie group a Poisson space which is the target of a momentum map for every Poisson action of the given group. Ortega and Ratiu [20], on the other hand, give preeminence to the symmetry/conservation relationship and, after fixing a particular action, find a space whose function algebra consists of the quantities which are conserved for all invariant hamiltonian systems. We now describe these two approaches. Details omitted here will appear in [5].

The construction by Evens and Lu may be seen as an extension of that of Chu [6]. (Also see [11].) In this earlier work, one has a Lie group G acting on a manifold P carrying a closed (but possibly degenerate) 2-form ω which is preserved by the action. For each x in P, we may pull back ω by the map $g \mapsto gx$ from G to P to get a closed, left-invariant 2-form $K(x)$ on G. Such forms correspond to 2-cocycles with constant real coefficients for the Lie algebra $\mathfrak{g}$; i.e. K is a (smooth) map from P to $Z^2(\mathfrak{g})$. The target $\mathfrak{g}^*$ of usual momentum maps is the same as the space $C^1(\mathfrak{g})$ of Lie algebra cochains, and the coboundary operator

is a natural map $\delta : \mathfrak{g}^* \to Z^2(\mathfrak{g})$. It turns out that $Z^2(\mathfrak{g})$ is the dual of a Lie algebra, and that δ is a Poisson map. If $J : P \to \mathfrak{g}^*$ is an equivariant momentum map in the usual sense, then the composition $\delta \circ J$ is equal to K. Notice that the map K exists and is canonically associated to every symplectic action, while J may not exist and is defined only up to an additive constant (which is in the kernel of δ). On the other hand, K may be much "cruder" than J. For instance, when $\mathfrak{g}$ is abelian, δ is the zero map, and so K is constant and equal to zero (i.e. the orbits of a hamiltonian action are isotropic), even though J may be nontrivial.

If a momentum map $J : P \to \mathfrak{g}^*$ is not equivariant, then $\delta \circ J$ differs from K by a constant element b of $Z^2(\mathfrak{g})$. If one adds the cocycle b to the Poisson structure on $\mathfrak{g}^*$ as a constant term, then J becomes a Poisson map and is equivariant with respect to an affine action of G on $\mathfrak{g}^*$ whose linear part is the coadjoint action. The translational part of this action is a 1-cocycle on G with values in $\mathfrak{g}^*$; the derivative of this cocycle at the identity of G may be identified with b. (See [22] or [23].) One may view $\mathfrak{g}^*$ with this affine Poisson structure as a Poisson submanifold of the dual of a central extension of $\mathfrak{g}$. Although one could use the 1-dimensional extension associated to b, it is interesting to use the universal central extension $X(g)$ of $\mathfrak{g}$ by $Z^2(\mathfrak{g})^*$ in which the bracket $[x, y]$ of two elements of $\mathfrak{g}$ is the linear functional on $Z^2(\mathfrak{g})$ given by $b \mapsto b(x, y)$. There is a natural Poisson map $\hat{\delta}$ from $X(\mathfrak{g})^*$ to $Z^2(\mathfrak{g})$. (It is not just the restriction map.) Any momentum map J for a symplectic action of G on P yields an equivariant Poisson map $\hat{J} : P \to X(\mathfrak{g})^*$ with $\hat{\delta} \circ \hat{J} = K$. Although $\hat{J}$ is not completely determined by the action, it does provide, for any action admitting a momentum map, a Poisson map which lifts K and which is as "refined" as J.

The definition of the map K is extended to Poisson actions of a Poisson Lie group G in the construction of Evens and Lu [10]. Their target is the space $\mathcal{L}(D\mathfrak{g})$ of *lagrangian subalgebras* of the *Drinfeld double* $D\mathfrak{g}$ of the Lie bialgebra $\mathfrak{g}$, the Lie algebra of the double group DG. The Drinfeld double is a Lie algebra structure on the vector space $\mathfrak{g} \oplus \mathfrak{g}^*$ together with the invariant inner product $((x, \xi), (y, \eta)) \mapsto \xi(y)\eta(x)$; a lagrangian subalgebra is a subalgebra which is maximal isotropic with respect to the inner product. (This is another example of a Dirac structure in a Courant algebroid.) Evens and Lu show that $\mathcal{L}(D\mathfrak{g})$ has the structure of an algebraic variety which is stratified by Poisson manifolds such that there is a dense subalgebra of continuous functions on $\mathcal{L}(D\mathfrak{g})$ whose restrictions to the strata are smooth and which is closed under the stratum-wise Poisson bracket. Given a Poisson action of G on P, they construct a natural equivariant map $EL : P \to \mathcal{L}(D\mathfrak{g})$ which is a smooth Poisson map into a single stratum. Using the theory of Dirac structures [8][14][15], it is possible to construct the equivariant map EL for arbitrary Poisson actions, though it is generally not continuous (though it is in fact smooth when P is symplectic). Nevertheless, EL can be shown to be a Poisson map in a certain sense. When G has the zero Poisson structure, the double $D\mathfrak{g}$ is the semidirect product of $\mathfrak{g}$ with $\mathfrak{g}^*$ via the coadjoint representation. Among the maximal isotropic subspaces of $\mathfrak{g} \oplus \mathfrak{g}^*$ are the graphs of antisymmetric mappings $\mathfrak{g} \to \mathfrak{g}^*$, which correspond to skew symmetric bilinear forms, or 2-cochains with real coefficients, on $\mathfrak{g}$. It turns out that the subspace is a subalgebra if and only if the cochain is a cocycle. (The corresponding subgroup is the graph of the $\mathfrak{g}^*$-valued cocycle on G connected with the affine action discussed above.) Thus, the target space $Z^2(\mathfrak{g})$ is contained in $\mathcal{L}(D\mathfrak{g})$ (as an open submanifold). The Poisson structure on $Z^2(\mathfrak{g})$ induced from the Evens-Lu structure on $\mathcal{L}(D\mathfrak{g})$ turns out to be identical to that which occurs in the construction of Chu, and the maps EL and K coincide when P is symplectic.

Continuing with the case where G has the zero Poisson structure, we also find among the lagrangian subalgebras of $D\mathfrak{g}$ the sums $\mathfrak{h} \oplus \mathfrak{h}^\perp$, where $\mathfrak{h}$ is an arbitrary subalgebra of $\mathfrak{g}$ and $\mathfrak{h}^\perp$ is its annihilator in $\mathfrak{g}^*$. These occur as the values of EL when the manifold P carries the zero Poisson structure : $EL(p) = \mathfrak{g}_p \oplus \mathfrak{g}_p^\perp$, where $\mathfrak{g}_p$ is the Lie algebra of the isotropy subgroup G_p of p in G. (The corresponding subgroup of the double group is the conormal bundle of G_p.) Thus, in the absence of any Poisson structure at all, we still have a "momentum map" which essentially assigns to each point in a G-manifold its isotropy submanifold. The level sets of this momentum map,or *isotropy types*, are invariant under any G-equivariant map.

Of course, there are no interesting G-invariant hamiltonian systems when P has the zero Poisson structure, since all hamiltonian vector fields are zero; nevertheless, this example provides a nice lead-in to the construction of Ortega and Ratiu [20], since their construction is motivated in part by the fact that conventional momentum maps may miss the "conservation of isotropy".

Ortega and Ratiu begin with a group G acting by automorphisms on a Poisson manifold M. Let $\mathcal{E}$ be the set of hamiltonian vector fields of G-invariant functions defined on G-invariant open subsets of M. The (local) flows of the vector fields in $\mathcal{E}$ form a pseudogroup $\mathcal{G}$ of Poisson diffeomorphisms between open subsets of M, and we may form the *momentum space* $M/\mathcal{G}$. In general, $M/\mathcal{G}$ is not a manifold, but it is a "smooth Poisson space" in the sense that, if we define the smooth functions on $M/\mathcal{G}$ to be the $\mathcal{G}$-invariant smooth functions on M, the space of these functions forms a Poisson algebra. Ortega and Ratiu's *optimal momentum map* is simply the natural projection map $\mathcal{J} : M \to M/\mathcal{G}$. Immediately from its definition, it has the property of being being conserved by the flow of any G-invariant hamiltonian. Since the elements of $\mathcal{G}$ commute with G, G acts in a natural way on the momentum space, and the optimal momentum map is equivariant. It is universal in the sense that, given any equivariant Poisson map $K : M \to P$ which is conserved by the flows of all G-invariant hamiltonians, there is a unique equivariant Poisson map $\phi : M/\mathcal{G} \to P$ such that $K = \phi \circ \mathcal{J}$. The optimal momentum map can be nontrivial even if a momentum map in the usual sense does not exist. For hamiltonian actions, it may be more refined than the usual momentum map J. For instance, in the case of a proper hamiltonian action, each level set of $\mathcal{J}$ is a connected component of the intersection of a level set of J with an isotropy type.

We remark that the optimal momentum map is very closely related to the Poisson Γ structure introduced by Condevaux, Dazord, and Molino [7] in a paper whose title is essentially the same as that of this one.

7 The Baer groupoid

We can forge closer links among some of the concepts in this paper with the aid of the *Baer groupoid* $\mathcal{B}G$ associated to a group G. The set of objects of this groupoid is the set $\mathcal{S}G$ of subgroups of G; a morphism from H to K is a subset of G which is at the same time a right coset Ka and a left coset bH. For such a morphism, it follows that $bH = aH$, so that $K = aHa^-1$, and K and H are conjugate. The product of two morphisms is the usual product of subsets in a group.

Thus, the $\mathcal{B}G$ orbits in $\mathcal{S}G$ are the conjugacy classes, and the isotropy subgroup $H \in \mathcal{S}G$

is isomorphic to the quotient NH/H, where NH is the normalizer of H in G. [4]

Any action of a group G on a space P induces an action of $\mathcal{B}G$ on P. The momentum map of this action assigns to each $x \in P$ the isotropy subgroup G_x, and the coset aG_x maps x to ax. The level sets of this momentum map are exactly the isotropy types, so we are close to the Evens–Lu momentum map. If the action is hamiltonian, we can extend the $\mathcal{B}G$ action on P to an action of the product groupoid $\mathcal{B}G \times T^*G$ with object space $\mathcal{S}G \times \mathfrak{g}^*$. The momentum map of this extended action has as level sets the intersection of the level sets of $J : P \to \mathfrak{g}^*$ with the isotropy types. The connected components of these level sets are, for proper actions, precisely the level sets of the optimal momentum map. Except for this question of connectedness, we have recovered the optimal momentum map using a natural construction involving a momentum space which depends only on the group and not on the particular (proper) action under consideration.

Further development of these ideas looks like a good project for the early 21st century.

[4] When G is a compact connected Lie group, the restriction of $\mathcal{B}G$ to the set of maximal tori is a transitive groupoid which might be called the *Weyl groupoid*, since its isotropy groups are the Weyl groups. A possible advantage of this groupoid over the individual Weyl groups is that it admits a nontrivial action of the automorphism group of G.

Bibliographie

[1] Alekseev, A., Malkin, A., and Meinrenken, E., *Lie group valued moment maps*, J. Differential Geom., **48** 1998, 445–495.

[2] A. Baer, *Zur Einführung des Scharbegriffs*, J. Reine Angew. Math., **160** 1929, 199-207.

[3] W. Borho and J-L. Brylinski, *Differential operators on homogeneous spaces II : Relative enveloping algebras*, Bull. Soc. Math. France, **117** 1989, 167–210.

[4] R. Brown, G. Danesh-Naruie and J. P. L. Hardy, *Topological groupoids : II. Covering morphisms and G-spaces*, Math. Nachr., **74** 1976, 143-156.

[5] H. Bursztyn and A. Weinstein,*Lectures on Momentum Maps* in preparation.

[6] B.Y. Chu, *Symplectic homogeneous spaces*, Trans. Amer. Math. Soc., **197** 1974, 154–159.

[7] M. Condevaux, P. Dazordand and P. Molino, *Géométrie du moment*, Travaux du Séminaire Sud-Rhodanien de Géométrie, I, Publ. Dép. Math. Nouvelle Sér. B, **88-1**, Univ. Claude-Bernard, Lyon, 1988, 131–160.

[8] T.J. Courant, *Dirac manifolds*, Trans. A.M.S., **319** 1990, 631-661.

[9] R. Dugas, *La Mécanique au XVII*^{e} *Siècle des Antécédents Scolastiques à la Pensée Classique*, Éditions du Griffon, Neuchatel, 1954.

[10] S. Evens and Lu, J-H., *On the variety of Lagrangian subalgebras*, preprint math.DG/9909005, *Ann. École Norm. Sup.*, to appear.

[11] V. Guillemin, *and S. Sternberg*, Symplectic Techniques in Physics, Cambridge Univ. Press, Cambridge, 1984.

[12] Y. Karshon, *Moment maps and non-compact cobordisms*, J. Diff. Geom., **49** 1998, 183–201.

[13] S. Lie, *Theorie der Transformationsgruppen*, Zweiter Abschnitt, unter Mitwirkung von Prof. Dr. Friedrich Engel, Teubner, Leipzig, 1890.

[14] Z. J. Liu, A. Weinstein and P. Xu, *Manin triples for Lie bialgebroids*, J. Diff. Geom, **45** 1997, 547-574.

[15] Z.-J. Liu, A. Weinstein and P. Xu, *Dirac structures and Poisson homogeneous spaces*, Commun. Math. Phys., **192** 1998, 121-144.

[16] J. H. Lu, *Momentum mappings and reduction of Poisson actions*, *Symplectic geometry, groupoids, and integrable systems, Séminaire sud-Rhodanien de géométrie à Berkeley 1989*, P. Dazord and A. Weinstein, eds., Springer-MSRI Series 1991, 209-226.

[17] J. Marsden and A. Weinstein, *Some comments on the history, theory, and applications of symplectic reduction*, dans *Quantization of Singular Symplectic Quotients*, N.P. Landsman, M. Pflaum, M. Schlichenmaier, eds., Birkhäuser, Basel, 2001, pp. 1-19.

[18] K. Mikami and A. Weinstein, *Moments and reduction for symplectic groupoid actions*, Publ. RIMS Kyoto Univ., **24** 1988, 121–140.

[19] E. Noether, *Invariante Variationsprobleme*, Nachr. v. d. Ges. d. Wiss. zu Göttingen, 1918, 235–257.

[20] J. P. Ortega, and T. S. Ratiu, *The optimal momentum map*, to appear in *Geometry, Dynamics, and Mechanics : 60th Birthday Volume for J.E. Marsden*, P. Holmes, P. Newton, and A. Weinstein, eds., Springer-Verlag, New York, 2002.

[21] S. Smale, *Topology and mechanics. I*, Invent. Math., **10** 1970, 305–331.

[22] J. M. Souriau, *Structure des Systèmes Dynamiques*, Dunod, Paris, 1970.

[23] J. M. Souriau, *Structure of dynamical systems. A symplectic view of physics*. Translated from the French by C. H. Cushman-de Vries. Translation edited and with a preface by R. H. Cushman and G. M. Tuynman., Birkhäuser, Boston, 1997.

[24] A. Weinstein, *Sophus Lie and symplectic geometry*, *Expos. Math.* **1** 1983, 95–96.

[25] A. Weinstein, *Coisotropic calculus and Poisson groupoids*, J. Math. Soc. Japan, **40** 1988, 705–727.

[26] A. Weinstein, *Poisson geometry of discrete series orbits, and momentum convexity for noncompact group actions*, Lett. Math. Phys., **56** 2001, 17-30.

Jets, transversalité, singularités : petite introduction aux grandes idées de René Thom

Marc Chaperon

Pour parler d'un monde aussi vaste en espérant être compris, j'ai construit cet exposé autour de l'exemple des équations différentielles multiformes, celui-là même que Thom avait choisi pour un exposé de "propagande" [16] au colloque Elie Cartan en 1971.

Présentant ainsi les choses par le petit bout de la lorgnette, je ne montrerai pas la profonde nécessité des jets[1] et de la transversalité en géométrie différentielle. En contrepartie, et c'est ma seule excuse, j'espère que l'honnête homme mathématique qui m'aura lu saura un peu de quoi il s'agit.

1 Position du problème

Soit l'équation différentielle

$$f(x, y, \frac{dy}{dx}) = 0, \tag{1.9}$$

où f est une fonction réelle C^∞ de trois variables réelles. Nous verrons qu'il est fructueux de l'écrire sous la forme du système

$$\begin{cases} f(x, y, z) = 0 \\ dy = z dx. \end{cases} \tag{1.10}$$

Une *solution* de l'équation différentielle (1.9) est une fonction réelle différentiable φ, définie sur un *intervalle* $I \subset \mathbb{R}$ ni vide, ni réduit à un point, et telle que

$$f\left(j^1\varphi(x)\right) = 0 \text{ pour tout } x \in I, \tag{1.11}$$

[1] J'aurais voulu mieux rendre justice à leur inventeur, Charles Ehresmann, un des très grands penseurs de la géométrie différentielle.

où $j^1\varphi$ est le *jet d'ordre* 1 de φ, défini par

$$j^1\varphi(x) := (x, \varphi(x), \varphi'(x)).$$

Lorsque φ est deux fois dérivable, il revient au même de dire que la courbe image $C = j^1\varphi(I)$, ensemble des $(x, y, z) \in \mathbb{R}^3$ avec $x \in I$, $y = \varphi(x)$ et $z = \varphi'(x)$, est une *solution généralisée* de (1.9), c'est-à-dire une *courbe intégrale* du système (1.10), ce qui signifie la chose suivante : la courbe C est contenue dans l'ensemble

$$S = f^{-1}(0)$$

des points (x, y, z) où f s'annule, et tout vecteur $(\delta x, \delta y, \delta z)$ tangent à C en un point (x, y, z) vérifie $\delta y = z\delta x$ (dans le cas d'une solution au sens habituel φ, cela exprime simplement que φ' est la dérivée de φ).

2 Étude géométrique de $S = f^{-1}(0)$

Au voisinage d'un point de S où la dérivée partielle $\Delta_z f$ n'est pas nulle, le théorème des fonctions implicites affirme que la surface S a pour équation $z = g(x, y)$, où la fonction implicite g est C^∞ comme f. L'équation (1.9) s'écrit donc localement comme une équation différentielle "ordinaire"

$$\frac{dy}{dx} = g(x, y), \tag{1.12}$$

dont les solutions géométriques sont les images des $j^1\varphi$ lorsque φ varie parmi les solutions φ de (1.12) au sens ordinaire.

Le sous-ensemble fermé Σ de $\mathbb{R}^3$ formé des points de S où $\Delta_z f$ s'annule joue donc un rôle-clé dans notre étude.

2.1 Génériquement, S est une surface C^∞

Hassler Whitney, un des premiers mathématiciens à prendre les fonctions C^∞ *en général* pour objet d'étude, est tombé sur des "monstres" comme le suivant :

Théorème 1 (Whitney) *Pour* tout *fermé F de $\mathbb{R}^n$, il existe une fonction $f \in C^\infty(\mathbb{R}^n, \mathbb{R})$ telle que $F = f^{-1}(0)$.*

Comme un fermé de $\mathbb{R}^n$ peut être très compliqué, même pour $n = 3$, il est légitime de souhaiter que S ne soit pas n'importe quel fermé de $\mathbb{R}^3$. Avant de justifier le titre de ce paragraphe, une petite

Définition 1 Un *point critique* d'une application $f \in C^\infty(M, N)$ est un point a de M où la différentielle de f n'est pas surjective ; on dit alors que $f(a)$ est une *valeur critique* de f. Les *valeurs régulières* de f sont les points de N qui ne sont pas valeurs critiques de f.

Pour toute valeur régulière b de f, l'ensemble $f^{-1}(b)$ des points de M où f prend la valeur b est une sous-variété de M, dont la codimension est la dimension de N.

Proposition 2.1 *"Presque toute" $f \in C^\infty(\mathbb{R}^3, \mathbb{R})$ admet 0 pour valeur régulière*[2]*, c'est-à-dire que sa différentielle $Df = (\Delta_x f, \Delta_y f, \Delta_z f)$ ne s'annule en aucun point du fermé $S = f^{-1}(0)$, qui est donc une sous-variété C^∞ de codimension 1 de $\mathbb{R}^3$, autrement dit une surface C^∞.*

Démonstration. Le théorème de Sard [10] nous dit que, pour chaque $f \in C^\infty(\mathbb{R}^3, \mathbb{R})$, presque tout $\varepsilon \in \mathbb{R}$ est une valeur régulière de f. En remplaçant f par $f - \varepsilon$, où ε est une telle valeur régulière de f, on obtient donc une fonction qui a 0 pour valeur régulière, si petit que soit ε. En élaborant ce résultat, Thom a obtenu le très important

Théorème 2 (lemme de transversalité dans les espaces de jets [12]) *Soient M, N deux variétés différentiables et W une sous-variété de l'espace $J^k(M, N)$ des jets d'ordre $k \in \mathfrak{N}$ d'applications de M dans N. Pour "presque toute" $f \in C^\infty(M, N)$, le jet d'ordre k de f est transverse à W.*

Expliquons les termes de cet énoncé dans le cas particulier où M, N sont des ouverts de $\mathbb{R}^m, \mathbb{R}^n$ respectivement. Le jet d'ordre k de f est alors l'application

$$j^k f : x \mapsto j^k_x f := (x, f(x), Df(x), D^2 f(x), \ldots, D^k f(x))$$

de M dans

$$J^k(M, N) := M \times N \times \prod_{j=1}^{k} L^j_s(\mathbb{R}^m, \mathbb{R}^n),$$

où $L^j_s(\mathbb{R}^m, \mathbb{R}^n)$ désigne l'espace vectoriel des applications j–linéaires symétriques de $(\mathbb{R}^m)^j$ dans $\mathbb{R}^n$ et $D^j f(x)$ la dérivée j–ième de f.[3]

La transversalité de $j^k f$ à W signifie la chose suivante : si l'intersection de W avec un ouvert U de $J^k(M, N)$ est l'ensemble des zéros d'une fonction $\pi \in C^\infty(U, \mathbb{R}^c)$ ayant zéro pour valeur régulière[4], alors 0 est valeur régulière de $(j^k f) \circ \pi$. Il en résulte immédiatement que si $j^k f$ est transverse à W, alors $(j^k f)^{-1}(W)$ est une sous-variété de M, de même codimension c que W.

Enfin, le "presque toute" s'entend au sens de la catégorie de Baire : l'ensemble des f telles que $j^k f$ soit transverse à W contient une intersection dénombrable d'ouverts denses de la topologie fine de Whitney, donc est lui-même dense[5] car c'est une topologie d'espace de Baire.

Fin de la preuve de la Proposition 2.1 Dire que 0 est valeur régulière de $f \in C^\infty(\mathbb{R}^3, \mathbb{R})$ s'exprime évidemment en disant qu'il n'existe aucun $a \in \mathbb{R}^3$ tel que $j^1_a f$ appartienne au sous-espace vectoriel W de $J^1(\mathbb{R}^3, \mathbb{R})$ formé des

$$j^1_{(x,y,z)}\varphi = (x, y, z, \varphi(x, y, z), \Delta_x\varphi(x, y, z), \Delta_y\varphi(x, y, z), \Delta_z\varphi(x, y, z))$$

[2]Cela se dit aussi, comme dans le titre de ce paragraphe : "Génériquement, f admet 0 pour valeur régulière".

[3]L'espace $\mathbb{R}^3$ où nous considérons (1.9) est donc $J^1(\mathbb{R}, \mathbb{R})$.

[4]Par définition d'une sous-variété, tout point de W est contenu dans un tel ouvert U.

[5]C'est un ensemble *résiduel*, ou, dans la terminologie répugnante de Bourbaki, *gras*.

avec $\pi(j^1_{(x,y,z)}\varphi) := (\varphi(x,y,z), \Delta_x\varphi(x,y,z), \Delta_y\varphi(x,y,z), \Delta_z\varphi(x,y,z)) = 0$ (en d'autres termes, W est l'ensemble des zéros de l'application linéaire surjective $\pi : J^1(\mathbb{R}^3,\mathbb{R}) \to \mathbb{R}^4$ ainsi définie). D'après le lemme de transversalité, pour presque toute $f \in C^\infty(\mathbb{R}^3,\mathbb{R})$, l'application $D(\pi \circ j^1 f)(a)$ est surjective en tout point a de $(\pi \circ j^1 f)^{-1}(0) = (j^1 f)^{-1}(W)$. Mais $D(\pi \circ j^1 f)(a)$ est une application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^4$, qui ne peut pas être surjective. Il n'existe donc aucun $a \in \mathbb{R}^3$ tel que $j^1 f(a)$ appartienne à W. ∎

Lorsque 0 est valeur régulière de f, la surface S a en tout point a un plan tangent T_aS, qui est le noyau de $df(a)$; l'ensemble Σ des points de S où $\Delta_z f$ s'annule est donc l'ensemble des points de S où ce plan tangent est *vertical.*

2.2 Génériquement, Σ est une courbe C^∞

Proposition 2.2 *Pour "presque toute" $f \in C^\infty(\mathbb{R}^3,\mathbb{R})$, le fermé Σ des points de $S = f^{-1}(0)$ où $\Delta_z f$ s'annule est une sous-variété C^∞ de codimension 2 de $\mathbb{R}^3$, autrement dit une courbe C^∞.*

Démonstration. Appliquons le lemme de transversalité de Thom au cas où $M = \mathbb{R}^3$, $N = \mathbb{R}$ et où W est le sous-espace vectoriel de codimension 2 de $J^1(\mathbb{R}^3,\mathbb{R})$ formé des

$$j^1_{(x,y,z)}\varphi = (x,y,z,\varphi(x,y,z), \Delta_x\varphi(x,y,z), \Delta_y\varphi(x,y,z), \Delta_z\varphi(x,y,z))$$

avec $\varphi(x,y,z) = \Delta_z\varphi(x,y,z) = 0$: pour presque toute $f \in C^\infty(\mathbb{R}^3,\mathbb{R})$, l'application $j^1 f$ est transverse à W, donc $\Sigma = (j^1 f)^{-1}(W)$ est bien une courbe C^∞. ∎

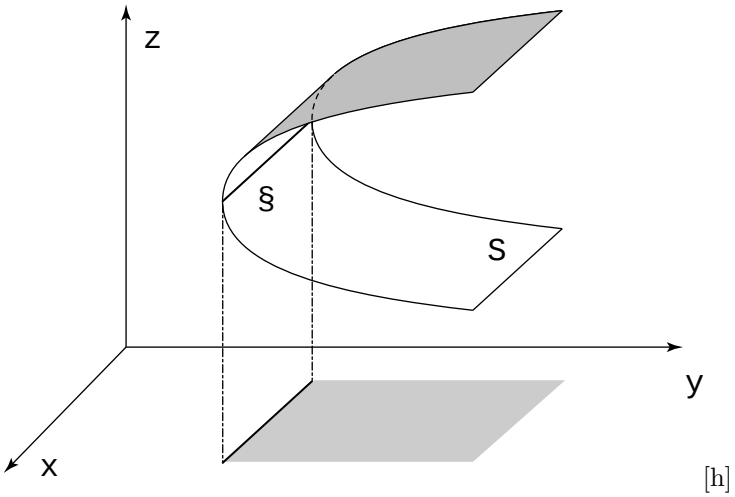
L'ensemble des points où Σ se projette mal sur le plan Oxy (c'est-à-dire que sa tangente est verticale) est alors

$$\Sigma_2 = (j^2 f)^{-1}(W_2), \text{ où } W_2 = \{j^2_p\varphi : \varphi(p) = \Delta_z\varphi(p) = \Delta_z^2\varphi(p) = 0\}.$$

Proposition 2.3 *Génériquement, $j^2 f$ est transverse à W_2, donc le fermé Σ_2 est formé de points isolés, dits* points-fronce *(les autres points de Σ sont les* points-plis*).* ∎

Le lieu singulier Σ de la projection de S sur le plan Oxy est baptisé "criminant" par V.I. Arnold[6], sa projection étant le *contour apparent* ou "discriminant".

[6] Ce qui est cohérent avec sa manière de parler des "criminels" responsables des programmes de mathématiques dans l'enseignement secondaire.



[h]

FIG. 5 – Situation générique près d'un point de $\Sigma \setminus \Sigma_2$ ("point-pli")

2.3 Formes normales

Proposition 2.4 *Au voisinage d'un point-pli (que l'on peut supposer être* 0 *en composant* f *avec une translation),*

$$f(x,y,z) = u(x,y,z)(z^2 + 2a(x,y)z + b(x,y)),$$

où u, a, b *sont des fonctions* C^∞ *vérifiant* $u(0) \neq 0$ *et* $a(0) = b(0) = 0$. *Dans les coordonnées locales*

$$(X,Y,Z) = (-b(x,y) + a(x,y)^2, y \text{ ou } x, z + a(x,y)),$$

la surface S *a pour équation* $X = Z^2$, *le "criminant"* Σ *est défini par* $X = Z = 0$, *et le discriminant, par* $X = 0$.

Proposition 2.5 *Près d'un point-fronce (que l'on peut supposer être* 0*),*

$$f(x,y,z) = u(x,y,z)(z^3 + 3a(x,y)z^2 + b(x,y)z + c(x,y)),$$

où u, a, b, c *sont* C^∞ *et vérifient* $u(0) \neq 0$ *et* $a(0) = b(0) = c(0) = 0$. *Dans les coordonnées locales*

$$(X,Y,Z) = ((-b + 3a^2)(x,y), (c + 2a^3 - ab)(x,y), z + a(x,y)),$$

la surface S *a pour équation* $Z^3 - XZ + Y = 0$, *le "criminant"* Σ *est défini par* $X = 3Z^2$, $Y = 2Z^3$, *et le discriminant, comme il se doit, par* $4X^3 = 27Y^2$.

Preuve La première partie de chaque proposition est un cas particulier du

Théorème 3 (théorème de préparation de Malgrange [6]) *Soit* f *un germe en* $0 \in \mathbb{R}^p \times \mathbb{R}$ *de fonction réelle* C^∞ *telle que* $\mathbb{R} \ni z \mapsto f(0,z)$ *soit nulle à l'ordre*

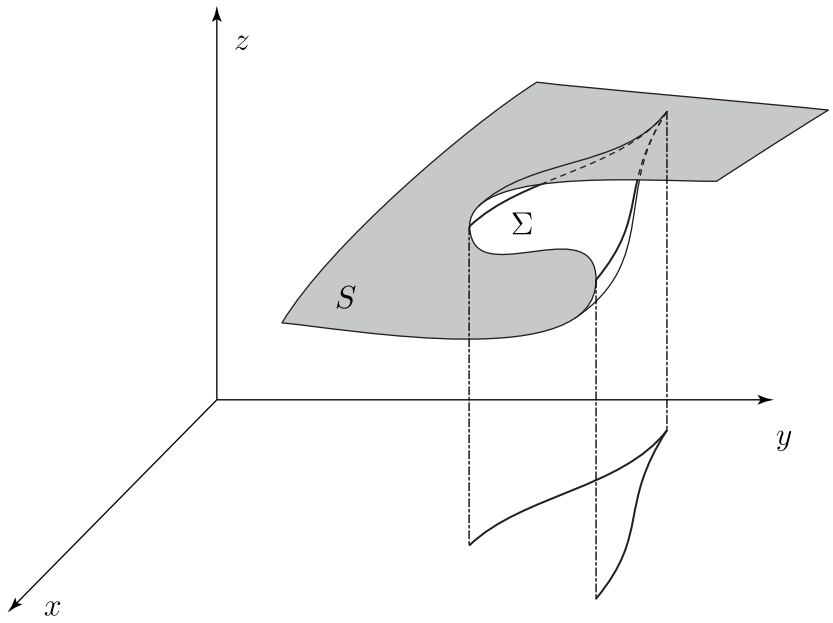


FIG. 6 – Situation générique près d'un point de Σ_2 ("point-fronce")

$k \in \mathfrak{N}$ *en* 0 (c'est-à-dire y soit nulle ainsi que ses k premières dérivées, la dérivée d'ordre $k+1$ étant non nulle). *On a alors*

$$f(\lambda, z) = u(\lambda, z)(z^{k+1} + a_1(\lambda)z^k + \cdots + a_k(\lambda)z + a_{k+1}(\lambda)),$$

où $u, a_1, \ldots, a_{k+1}$ *sont* C^∞ *et vérifient* $u(0) \neq 0$ *et* $a_1(0) = \cdots = a_{k+1}(0) = 0$.

Il reste à vérifier que X, Y, Z forment bien un système de coordonnées locales. Dans le cas du pli, cela vient de ce que $df(0) \neq 0$ signifie $db(0) \neq 0$, c'est-à-dire $dX(0) \neq 0$. On choisit donc $Y = y$ pour $\frac{\Delta X}{\Delta x}(0) \neq 0$ et $Y = x$ sinon. Dans le cas de la fronce, cela vient de ce que la transversalité de $j^2 f$ à W_2 se traduit par l'indépendance linéaire de $df(0)$, $d\Delta_z f(0)$ et $d\Delta_z^2 f(0)$, c'est-à-dire de $da(0)$ et $db(0)$ et donc de $dX(0)$ et $dY(0)$.

Pour conclure, il ne reste plus qu'à remarquer que le changement de variables préserve la notion de verticale, puisque X et Y dépendent seulement de x et y. ∎

Le théorème de préparation (dû à Weierstrass dans le cadre analytique, l'énoncé C^∞ étant plus difficile parce qu'il n'y a pas unicité) est un des outils essentiels de la théorie des singularités ; au modeste niveau de cet exposé, les formes normales du pli et surtout de la fronce ne sont pas si faciles à obtenir sans lui.

3 "Digression" au cœur du sujet

3.1 Conflits et bifurcations

Oubliant temporairement la structure de contact $dy = zdx$, supposons que f "dérive d'un potentiel" V par rapport à z, $f(x, y, z) = \Delta_z V(x, y, z)$ (ce qui est toujours le cas) et que ce soit ce potentiel qui nous intéresse. Au dessin représentant la surface S d'équation $\Delta_z V = 0$

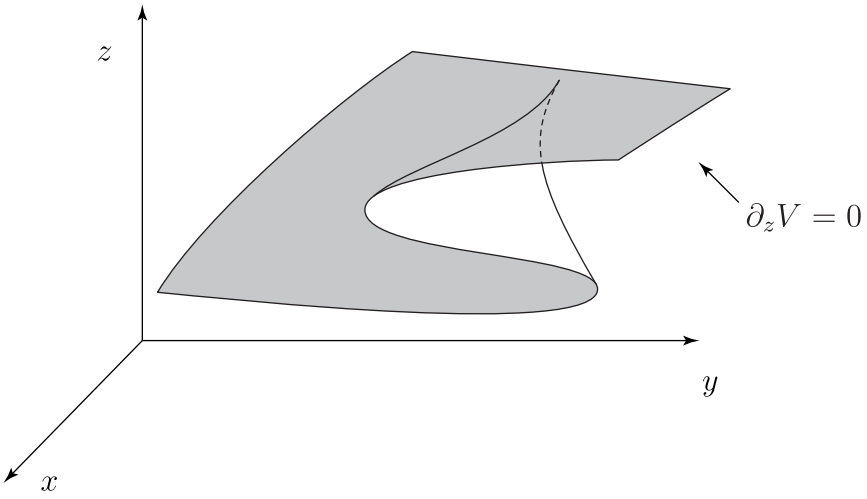
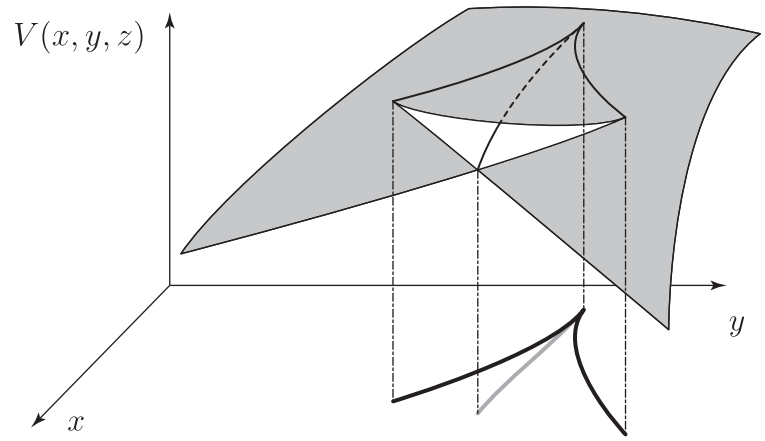


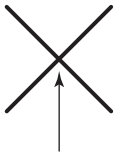
FIG. 7 – Situation générique près d'un point de Σ_2 ("point-fronce")

est associé le "diagramme de Cerf" ou "front d'onde" formé de tous les $(x, y, V(x, y, z))$ avec $\Delta_z V(x, y, z) = 0$:

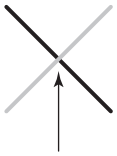
Au discriminant (en gras sur la figure), ensemble des (x, y) tels que $V_{x,y}$ ait au moins un point critique dégénéré, s'ajoute donc l'ensemble des (x, y) tels que la fonction $V_{x,y} : z \mapsto V(x, y, z)$ prenne la même valeur en deux points critiques disjoints, c'est-à-dire la courbe (en grisé sur la figure) projection des auto-intersections du front d'onde.

Génériquement, il y a donc dans le plan Oxy des paramètres une courbe en gras ("strates de bifurcation" dans la terminologie de R. Thom) et une courbe grisée ("strates de conflit"), dont voici les autres singularités :

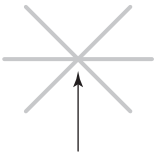




$V_{x,y}$ a deux points critiques dégénérés



$V_{x,y}$ a deux points critiques non dégénérés au même niveau et un point critique dégénéré



$V_{x,y}$ a trois points critiques non dégénérés au même niveau.

Stratification du plan Oxy des paramètres Les (x, y) qui ne sont sur aucune de ces deux courbes sont ceux pour lesquels $V_{x,y}$ est une *fonction de Morse excellente*, c'est-à-dire qu'elle a des points critiques tous non-dégénérés et y prend des valeurs toutes différentes. Lorsqu'on s'intéresse à des familles génériques de potentiels périodiques par rapport à z (par exemple de période 1, chaque $V_{x,y}$ étant donc vu comme une fonction C^∞ sur le tore $\mathbb{T} = \mathbb{R}/\mathbb{Z}$)[7], ces (x, y) forment un *ouvert* (évidemment dense) U_0. Son complémentaire admet naturellement une *stratification*, c'est-à-dire une partition en sous-variétés deux à deux disjointes : des strates de dimension 0, qui sont les projections des points-fronce et les autres intersections entre les courbes que nous venons de décrire ; des strates de dimension 1, formées de ce qui reste des courbes.

Cette stratification possède d'excellentes propriétés ("propriétés A et B de Whitney"). Introduites par Whitney [18], les stratifications jouent un rôle essentiel en théorie de singularités [14, 8, 4].

Si u et v sont dans une même composante connexe de la "grande strate" ouverte U_0, alors V_u et V_v sont *équivalents* : il existe des difféomorphismes (changements de variables)

[7] Bien entendu, la périodicité exclut qu'elle soit excellente comme fonction sur $\mathbb{R}$.

$\varphi : \mathbb{T} \to \mathbb{T}$ et $\psi : \mathbb{R} \to \mathbb{R}$ de classe C^∞ tels que $V_v = \psi^{-1} \circ V_u \circ \varphi$. Cela résulte du théorème plus précis suivant, appliqué à $f_t = V_{c(t)}$, où $[0,1] \ni t \mapsto c(t)$ est un chemin C^∞ joignant u à v dans U_0 :

Proposition 3.1 (stabilité des fonctions excellentes sur $\mathbb{T}$) *Pour toute fonction réelle $[0,1] \times \mathbb{T} \ni (t,x) \mapsto f_t(x)$ de classe C^∞ telle que chaque f_t soit une fonction excellente sur $\mathbb{T}$, il existe deux isotopies*[8] *(φ_t) et (ψ_t) de classe C^∞ de $\mathbb{T}$ et $\mathbb{R}$ respectivement telles que $\psi_t^{-1} \circ f_t \circ \varphi_t = f_0$ pour tout $t \in [0,1]$.*

Nous détaillons en appendice la démonstration de cette proposition parce qu'elle donne, dans un cas élémentaire, un premier aperçu substantiel des idées à l'œuvre dans les travaux de Thom et Mather sur les singularités.

3.2 Multijets

Pour montrer que les singularités génériques des familles à deux paramètres x, y de fonctions d'une variable z sont bien telles que nous les avons décrites, on applique le lemme de transversalité aux multijets [7] : pour tout entier $s > 0$, le s–jet d'ordre k de $\varphi : M \to N$ est

$$ {}_sj^k\varphi : (z_1, \ldots, z_s) \mapsto (j^k\varphi(z_1), \ldots, j^k\varphi(z_s)), $$

défini sur l'ouvert ${}_sM$ de M^s formé des $(z_1, \ldots, z_s)$ avec $z_1, \ldots, z_s$ tous différents.

Ainsi, dans notre exemple, les intersections conflit-bifurcation peuvent se traiter en remarquant que, pour presque toute $V \in C^\infty(\mathbb{R}^3, \mathbb{R})$, l'application

$$ (x, y, (z_1, z_2, z_3)) \mapsto {}_3j^2V_{x,y}(z_1, z_2, z_3) $$

est transverse à la sous-variété formée des $(j^2\varphi(z_1), j^2\varphi(z_2), j^2\varphi(z_3))$ avec $\varphi(z_1) - \varphi(z_2) = \varphi'(z_1) = \varphi'(z_2) = 0$ (conflit) et $\varphi'(z_3) = \varphi''(z_3) = 0$ (bifurcation).

3.3 La stratification naturelle des espaces de fonctions

La généricité de la transversalité dans ce qui précède vient, en dernière analyse[9], de ce que chacune des applications

$$ (\varphi, (z_1, \ldots, z_s)) \mapsto {}_sj^k\varphi(z_1, \ldots, z_s) $$

est une "submersion" de $C^\infty(\mathbb{R}, \mathbb{R}) \times {}_s\mathbb{R}$ dans ${}_sJ^k(\mathbb{R}, \mathbb{R})$ (on peut enlever les guillemets pour l'application analogue de $C^r(\mathbb{T}, \mathbb{R}) \times {}_s\mathbb{T}$ dans ${}_sJ^k(\mathbb{T}, \mathbb{R})$, $k < r < \infty$, où $\mathbb{T}$ est le cercle $\mathbb{R}/\mathbb{Z}$).

On peut donc imaginer que la modeste stratification du plan Oxy dessinée ci-dessus pour un V générique est une coupe bidimensionnelle d'une stratification analogue de $C^\infty(\mathbb{T}, \mathbb{R})$ se traduisant algébriquement en termes de multijets. Cette vision de R. Thom [15], précisée entre autres par J.N. Mather [8, 4], a servi de cadre conceptuel à de très importants travaux de Jean Cerf [2, 3] ou, plus récemment, de Victor Vassiliev [17] en topologie différentielle.

[8] Une *isotopie* C^∞ d'une variété M est un chemin $[0,1] \ni t \mapsto \varphi_t$ dans le groupe des difféomorphismes de M sur elle-même tel que $(t,x) \mapsto \varphi_t(x)$ soit C^∞ et que $\varphi_0 = \mathrm{Id}_M$.

[9] C'est l'essence de la preuve du lemme de transversalité.

Donnons par exemple une idée de la démonstration du théorème du h–cobordisme dans [2] (voir aussi [9]) : soit W un cobordisme entre deux variétés compactes (sans bord) V_0 et V_1, c'est-à-dire une variété compacte à bord dont le bord ΔW est la réunion disjointe de V_0 et V_1 ; on suppose que c'est un *h–cobordisme,* ce qui signifie que les injections $V_i \hookrightarrow W$ sont des équivalences d'homotopie. Le théorème du h–cobordisme, dû à Smale, affirme que *si W est simplement connexe et de dimension au moins* 6*, alors c'est un cobordisme* trivial *: il existe un difféomorphisme g de W sur $V_0 \times [0,1]$* (en particulier, V_0 et V_1 sont difféomorphes).

La démonstration repose sur une remarque très simple : pour que W soit un cobordisme trivial, il faut et il suffit qu'il existe une fonction *sans point critique* $f : W \to [0,1]$ prenant la valeur i en tout point de V_i $(i = 0, 1)$: une métrique riemannienne étant fixée sur W, on peut alors prendre pour difféomorphisme g^{-1} l'application qui à $(x_0, t) \in V_0 \times [0,1]$ associe le point de la ligne de gradient issue de x_0 où f prend la valeur t. Cerf considère donc l'espace $\mathcal{F}$ des fonctions $\varphi \in C^\infty(W, \mathbb{R})$ valant i sur V_i $(i = 0, 1)$, y "tire au hasard" une fonction f_0, qui est (selon toute probabilité…) une fonction de Morse excellente, et construit un chemin $[0,1] \ni t \mapsto f_t$ dans $\mathcal{F}$ d'origine f_0 tel que $(t, x) \mapsto f_t(x)$ soit C^∞ et f_1 sans point critique. Pour cela, le chemin, issu de la "grande strate" (ouverte) formée des fonctions excellentes va devoir traverser la strate de codimension 1 de la stratification naturelle, chaque traversée ayant pour effet de changer l'ordre des valeurs critiques ou de faire naître ou mourir (par couples) des points critiques, jusqu'à ce qu'il n'y en ait plus.

La même vision est (plus explicitement) à l'œuvre dans [3] et dans les travaux récents de Vassiliev [17] sur les nœuds, un nœud étant un élément de la "grande strate" (ouverte) de $C^\infty(\mathbb{T}, \mathbb{R}^3)$, formée des immersions injectives.

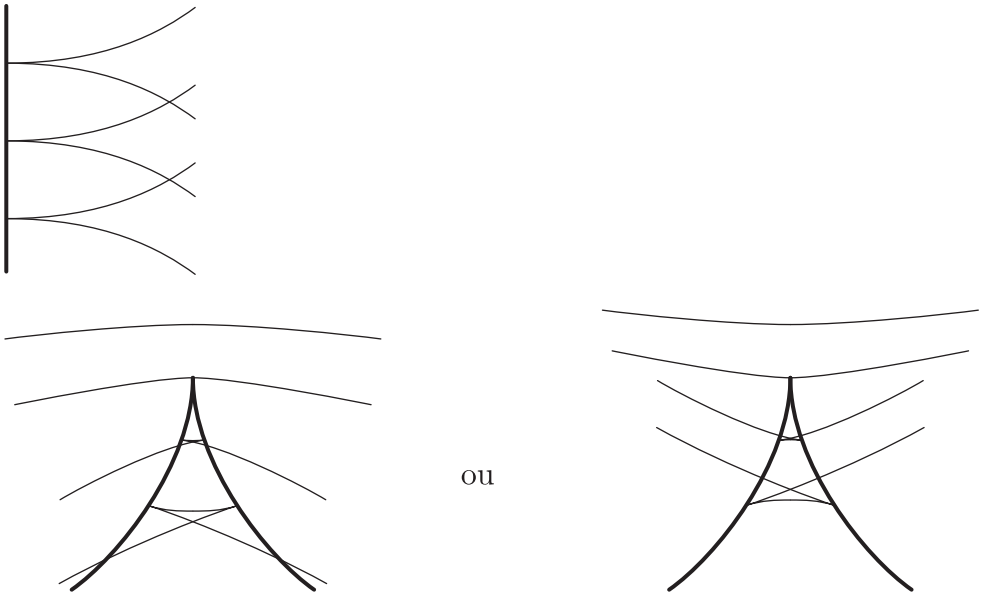
4 Retour à l'équation (1.9)

Nous allons en fait, pour une fonction f générique, étudier le système (1.10). En tout point $p = (x, y, z)$ de la surface $S = f^{-1}(0)$, il y a deux plans intéressants : le plan tangent T_pS, noyau de $df(p)$, et le plan de contact $\mathcal{K}_p$ d'équation $dy = z\,dx$. Ils coïncident sur un fermé Σ_0 de S, génériquement formé de points-plis isolés p (pour ne pas s'en étonner, il faut remarquer que chaque $\mathcal{K}_p$ est vertical). Le champ de droites $S \setminus \Sigma_0 \ni p \mapsto \chi_p = T_pS \cap \mathcal{K}_p$ s'intègre en un feuilletage de $S \setminus \Sigma_0$ par des courbes, appelées *caractéristiques* de S ou *solutions généralisées de* (1.9).

Chacune d'entre elles, au voisinage d'un point p où sa tangente χ_p n'est pas verticale (ce qui signifie que $p \in S \setminus \Sigma$), se projette sur le graphe d'une solution au sens ordinaire. Au voisinage d'un point-pli, elles se projettent ainsi :

Au voisinage d'un point-fronce, ainsi :

Un théorème de D. Bennequin affirme qu'aucune de ces solutions généralisées n'est une courbe fermée entourant un disque dans S ; on en déduit [1] l'existence de structures de contact "exotiques" sur $\mathbb{R}^3$.



4.1 La stratégie de Lie

Si l'on oublie la projection $J^1(\mathbb{R},\mathbb{R}) \ni j^1_x\varphi \mapsto j^0_x\varphi \in J^0(\mathbb{R},\mathbb{R})$, c'est-à-dire $(x,y,z) \mapsto (x,y)$, et qu'on ne retient que la structure de contact $dy = zdx$, les points $a \in \Sigma \setminus \Sigma_0$ n'ont rien d'exceptionnel : on peut ramener $a = (x_0, y_0, z_0)$ en 0 par la transformation

$$(x,y,z) \mapsto (x - x_0, y - y_0 - z_0(x - x_0), z - z_0),$$

qui conserve la forme de contact $dy - zdx$ et donc la structure de contact (on dit que c'est une *transformation de contact*) ; puisque désormais $a = 0$, la condition $T_aS \neq \mathcal{K}_a$ s'écrit $\Delta_x f(0) \neq 0$, donc S admet au voisinage de 0 l'équation

$$x = g(y,z), \text{ avec } g \in C^\infty(\mathbb{R}^2,\mathbb{R}).$$

Une autre transformation de contact, la *transformation de Legendre*

$$(x,y,z) \mapsto (-z, y - zx, x)$$

transforme alors cette équation, près de 0, en une équation différentielle "ordinaire" $\frac{dy}{dx} = G(x,y)$, dont les solutions $y = \varphi(x)$, identifiées chacune à la courbe image de $j^1\varphi$, sont les images des solutions géométriques locales de (1.9) au voisinage de a par les deux transformations de contact précédentes.

Devant "l'excès de structure" de $J^1(\mathbb{R},\mathbb{R})$, doté à la fois d'une structure de contact et d'une projection sur $J^0(\mathbb{R},\mathbb{R})$, la stratégie consistant à oublier dans un premier temps la projection est celle adoptée par Sophus Lie dans ses travaux sur les équations aux dérivées partielles.

4.2 La stratégie de Gromov

Le *h–principe* (ou *principe d'homotopie*) est un énoncé très général que Gromov [5] abstrait entre autres de ce qu'avait fait Smale dans sa thèse sur les immersions. La stratégie est exactement inverse de celle de Lie[10], puisqu'on commence par oublier la structure de contact en se concentrant sur la projection.

Par exemple, la démonstration de Smale se raconte ainsi : dans l'espace $J^1(M,\mathbb{R}^N)$, on considère l'ouvert U formé des jets d'immersions, c'est-à-dire des $j^1_x\varphi$ avec $d\varphi_x$ injective. Le h–principe dans ce cas dit que toute section de la projection $J^1(M,\mathbb{R}^N) \ni j^1_x\varphi \mapsto x \in M$ qui prend toutes ses valeurs dans U est homotope *parmi les sections à valeurs dans* U à une section holonome, c'est-à-dire au jet d'ordre 1 d'une immersion[11].

Parmi les autres idées du livre foisonnant [5] de Gromov figure, sous l'appellation d'*intégration convexe*, une abstraction de la preuve par Nash et Kuiper[12] du fascinant théorème de plongement isométrique C^1, qui dit par exemple que *pour qu'une variéte riemannienne compacte admette un plongement isométrique* C^1 *dans* $\mathbb{R}^N$*, il faut et il suffit qu'elle puisse s'y plonger.* Un des problèmes que posent de tels plongements isométriques est d'être rebelles, faute d'une différentiabilité suffisante, aux idées de transversalité qui ont fait en partie le succès de Thom : ce n'est pas en ces termes que l'on peut dire quelles sont les singularités génériques d'une feuille de papier qui se froisse...

5 Appendice : preuve de la proposition 3.1

Nous allons appliquer la *méthode du chemin*, à la base de toute la théorie des singularités : l'idée est qu'une isotopie (g_t) s'obtient en intégrant un champ de vecteurs dépendant du temps $(X_t)_{0\le t\le 1}$, donné (forcément) par $X_t = (\frac{d}{dt}g_t)\circ g_t^{-1}$, lequel la détermine puisque $g_t(x)$ est la valeur au temps t de la solution de l'équation différentielle $\frac{dy}{dt} = X_t(y)$ qui vaut x au temps 0 (ici intervient le fait que g_0 est l'identité).

Inversement, un champ de vecteurs dépendant du temps, lorsque toutes les solutions de l'équation différentielle associée existent sur $[0,1]$ tout entier, engendre une isotopie. Nous allons donc chercher des champs de vecteurs (c'est-à-dire des fonctions réelles) dépendant du temps (X_t) sur $\mathbb{T}$ et (Y_t) sur $\mathbb{R}$, engendrant des isotopies (φ_t) et (ψ_t) telles que $\psi_t^{-1}\circ f_t\circ\varphi_t = f_0$ pour tout $t\in[0,1]$. Comme cette relation est automatiquement vérifiée pour $t=0$, il est équivalent de demander que $\frac{d}{dt}(\psi_t^{-1}\circ f_t\circ\varphi_t) = 0$ pour tout t. Or, puisqu'en dérivant par rapport à t la relation $\psi_t^{-1}\circ\psi_t = \mathrm{Id}$ on obtient $\frac{d}{dt}(\psi_t^{-1}) = -(\psi_t^{-1})'\cdot Y_t$, on voit que

$$\begin{aligned}\tfrac{d}{dt}(\psi_t^{-1}\circ f_t\circ\varphi_t) &= (\tfrac{d}{dt}(\psi_t^{-1}))\circ f_t\circ\varphi_t + \big((\psi_t^{-1})'\circ f_t\circ\varphi_t\big)\cdot(\tfrac{d}{dt}f_t)\circ\varphi_t\\ &\quad + \big((\psi_t^{-1})'\circ f_t\circ\varphi_t\big)\cdot((f_t)'\circ\varphi_t)\cdot(\tfrac{d}{dt}\varphi_t)\\ &= \big(\big((\psi_t^{-1})'\circ f_t\big)\cdot\big(-Y_t\circ f_t + (\tfrac{d}{dt}f_t) + (f_t)'\cdot X_t\big)\big)\circ\varphi_t,\end{aligned}$$

[10] Mais il ne s'agit pas de gagner les mêmes batailles, le h–principe concernant des systèmes moins contraints que ceux auxquels s'applique le point de vue de Lie.

[11] L'exposé au séminaire Bourbaki [11] où Thom dégage remarquablement les idées de cette démonstration constitue un maillon important entre Smale et Gromov.

[12] Gromov a traduit Nash en russe.

qui est identiquement nul si et seulement si

$$-Y_t(f_t(x)) + \tfrac{d}{dt} f_t(x) + (f_t)'(x) \cdot X_t(x) = 0 \text{ pour tout } (t, x).$$

Si l'on veut pouvoir diviser par $(f_t)'(x)$ pour obtenir $X_t(x)$, il faut au préalable choisir les valeurs que prend Y_t sur les valeurs critiques de f_t de manière que $Y_t(f_t(x)) = \frac{d}{dt} f_t(x)$ pour tout point critique x de f_t. Un tel choix de Y_t est possible (ici intervient le fait que les valeurs de f_t en ses points critiques sont toutes distinctes) ; si on l'effectue de manière "canonique", par exemple en choisissant le polynôme d'interpolation de Lagrange , on obtient une fonction $(t, x) \mapsto Y_t(x)$ de classe C^∞, toutes les fonctions f_t ayant (par compacité de $\mathbb{T}$ et parce qu'elles sont toutes de Morse) le même nombre fini de points critiques, dont chacun est une fonction C^∞ de t. En multipliant tous les Y_t ainsi obtenus par une même fonction C^∞ réelle d'une variable réelle à support compact, égale à 1 au voisinage de toutes les valeurs critiques des f_t, on obtient un champ de vecteurs dépendant du temps (Y_t) sur $\mathbb{R}$, possédant les propriétés requises et *à support compact*, ce qui garantit qu'il engendre une isotopie (ψ_t).

La formule $X_t = \frac{1}{(f_t)'}(Y_t \circ f_t - \frac{d}{dt} f_t)$ définit alors un champ de vecteurs (X_t) dépendant du temps *de classe* C^∞ sur $\mathbb{T}$, comme on le voit en appliquant à X_t et $(f_t)'$ la formule de la moyenne au voisinage de chaque point critique de f_t (ici intervient le fait que ceux-ci sont non dégénérés). Comme $\mathbb{T}$ est compact, (X_t) engendre une isotopie (φ_t) et nous en avons terminé puisque, par construction, $\psi_t^{-1} \circ f_t \circ \varphi_t = f_0$ pour tout $t \in [0, 1]$. ∎

La possibilité de diviser par $(f_t)'$ est évidemment propre aux fonctions réelles d'une variable réelle, mais le reste est typique de la théorie des singularités. Pour reprendre le titre du premier article de [7], nous avons montré dans un cas particulier que *la stabilité infinitésimale entraîne la stabilité.*

Bibliographie

[1] D. Bennequin, *Entrelacements et équations de Pfaff.* III^{e} rencontre de géométrie du Schnepfenried, Volume 1, Astérisque **107–108** (1983), 87–161.

[2] J. Cerf et A. Gramain, *Le théorème du h–cobordisme (Smale).* Cours professé au printemps 1966 à la Faculté des Sciences d'Orsay. Secrétariat Mathématique de l'École Normale Supérieure, 1968.

[3] J. Cerf, *La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie.* Publ. IHES **39** (1970), 5–173.

[4] A. Chenciner, *Travaux de Thom et Mather sur la stabilité topologique.* Séminaire Bourbaki, 25ème année (1972/73), exposé 424, Springer Lecture notes in Math. **383** (1974), 116–140.

[5] M. Gromov, *Partial Differential Relations.* Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer-Verlag, Berlin, 1986.

[6] B. Malgrange, *Le théorème de préparation en géométrie différentielle.* Séminaire H. Cartan, 1962/63, exposés 11, 12, 13, 22.

[7] J.N. Mather, *Stability of C^{∞} mappings,*
I. Annals of Math. **82** (1968), 89–109,
II. Annals of Math. **89** (1969), 259–291,
III. Publ. IHES **35** (1969), 127–156,
IV. Publ. IHES **37** (1970), 223–248,
V. Advances in Math. **4** (1970), 301–335,
VI. Proc. Singularities Symp. Liverpool 1971, Springer Lecture notes in Math. **192** (1971), 207–255.

[8] J.N. Mather, *How to stratify mappings and jet spaces.* Singularités d'applications différentiables, Plans-sur-Bex 1975. Springer Lecture Notes in Mathematics **535** (1976), 128–176.

[9] J.W. Milnor, *Lectures on the h–cobordism theorem.* Princeton University Press, 1968 ?.

[10] A. Sard, *The measure of the critical points of differentiable maps,* Bull. Amer. Math. Soc. **48** (1942), 883–890.

[11] René Thom, *La classification des immersions.* Séminaire Bourbaki 1957/1958, Exposé 157, 11 pp., Secrétariat mathématique, Paris 1958.

[12] René Thom, *Un lemme sur les applications différentiables,* Bol. Soc. Mat. Mexicana (1956), 59–71.

[13] René Thom, *Local topological properties of differentiable mappings.* Differential Analysis, Bombay Colloq. (1964) Oxford Univ. Press, 191–202.

[14] René Thom, *Ensembles et morphismes stratifiés,* Bull. Amer. Math. Soc. **75** (1969), 240–284.

[15] René Thom, *The bifurcation subset of a space of maps,* in *Manifolds,* Amsterdam, 1970, Springer Lecture Notes in Mathematics , 202–208.

[16] René Thom, *Sur les équations différentielles multiformes et leurs intégrales singulières,* Bol. Soc. Mat. Bras. **3** (1972), 1–11.

[17] V.A. Vassiliev, *Cohomology of knot spaces,* Theory of singularities and its applications, Adv. Soviet Math. **1** (1990), publié par l'AMS, 23–69.

[18] H. Whitney, *Tangents to an analytic variety,* Ann. of Math. **81** (1965), 496–549.

Naissance des feuilletages, D'Ehresmann-Reeb à Novikov

André Haefliger[1]

Cet exposé est centré sur la thèse de Reeb et sur son influence sur les premiers développements de la théorie des feuilletages à travers les problèmes qui y sont abordés.

Après avoir rappelé dans un premier paragraphe la notion de complète intégrabilité pour un champ de p-plans, j'évoque dans le deuxième paragraphe la naissance de la théorie des feuilletages en examinant dans l'ordre chronologique les notes aux Comptes Rendus de l'Académie des Sciences de Paris publiées par Ehresmann et Reeb de 1944 à 1948, année de la soutenance de la thèse de Reeb. Entre 1948 et 1954 Ehresmann y apportera des compléments importants, en particulier la notion d'holonomie, que nous décrivons dans le troisième paragraphe. Dans le paragraphe suivant, je mentionne les quelques contributions de ma thèse (55-58) et l'apport fondamental de Novikov en 1964-65 qui marque le début de l'âge d'or de la théorie.

Pour terminer, je mentionne 5 problèmes qui s'imposaient à tout lecteur de la thèse de Reeb au début des années cinquante. Ces problèmes ont joué un rôle considérable dans le développement de la théorie et j'indique brièvement les principales étapes de leur résolution.

Je remercie E. Ghys, A. Phillips et J. Pradines pour leurs commentaires constructifs.

1 Préliminaires

Commençons par rappeler quelques notions mathématiques bien connues déjà au 19ème siècle[2], tout au moins dans un ouvert de $\mathbb{R}^n$.

Un champ K d'éléments de contact de dimension p (ou pour simplifier un champ de p-plans) de classe C^r sur une variété V est la donnée pour chaque point x de V d'un sous-espace $K(x)$ de l'espace tangent à V en x, variant r-différentiablement avec x. Si n est la dimension de V, on dit que $q = n - p$ est la codimension de K. Localement un tel champ peut être donné par l'annulation de p formes différentielles de degré 1 (formes de Pfaff) linéairement indépendantes en chaque point. Une sous-variété intégrale (sous-entendu de dimension p)

[1]Section de Mathématiques, 2-4 Rue du Lièvre, 1211 Genève 24

[2]On pourra consulter à ce propos l'article *Propriétés générales des systèmes d'equations aux dérivées partielles. Equations linéaires du premier ordre*, de G. Floquet d'après E. von Weber, paru dans l'Encyclopédie des Sciences mathématiques, (1910), Tome II, vol.4, p. 1-55.

pour un tel champ est une sous-variété connexe de V de dimension p tangente en chacun de ses points x à $K(x)$. Le champ K est dit complètement intégrable si par chaque point de V passe une sous-variété intégrale. Il en résulte (tout au moins si $r \geq 1$) qu'il existe au voisinage de chaque point de V des coordonnées locales $(x_1, \ldots, x_n)$ telles que K soit défini dans ces coordonnées par l'annulation des q formes $dx_{n-p+1}, \ldots, dx_n$, les varétés intégrales dans ce système de coordonnées étant des morceaux ouverts des p-plans parallèles aux p premiers axes de coordonnées.

Lorsque K est complètement intégrable, par chaque point de V passe une variété intégrale complète (i.e. maximale) unique qui est une sous-variété immergée et V est l'union disjointe des variétés intégrales complètes.

Si $p = 1$, tout champ K de 1-plans de classe C^1 est complètement intégrable (cela résulte de l'existence de solutions pour les équations différentielles). Ce n'est plus le cas si $p > 1$. Par exemple dans l'espace euclidien à 3 dimensions, le champ de 2-plans donnés par l'annulation de la forme de Pfaff

$$\omega = dz - xdy + ydx$$

n'est pas complètement intégrable. Le long de l'axe Oz le champ est horizontal et l'intersection avec un cylindre d'axe Oz est un champ de directions tangentes à des spirales.

D'une manière générale, pour un champ de plans de codimension 1 donné par une forme de Pfaff ω, la condition d'intégrabilité équivaut à $\omega \wedge d\omega = 0$. Dans ce cas il existe localement une fonction non nulle λ, appelée un *facteur intégrant*, et une fonction ϕ, appelée une *intégrale première*, telles que $\omega = \lambda d\phi$. Les variétés de niveau de ϕ sont les variétés intégrales. Carathéodory [5] a donné en 1909 une caractérisation géométrique locale de la complète intégrabilité d'une forme de Pfaff ω, à savoir : ω est complètement intégrable si et seulement si, pour tout voisinage U de tout point x, il existe un point de U qui ne peut être relié à x par une courbe dans U tangente au noyau de ω. Il a utilisé cette caractérisation pour exprimer de manière remarquablement concise et conceptuelle le second principe de la thermodynamique[3]

Tout champ de vecteurs partout non nul définit un champ de 1-plans dont les variétés intégrales complètes sont les trajectoires ou solutions maximales de l'équation différentielle correspondante. Poincaré [2] a initié à la fin du 19-ème siècle l'étude qualitative globale des trajectoires de ce type, suivie par un nombre considérable de travaux, en particulier dans le cas où la variété ambiante est une surface (travaux de Bendixson [4], Denjoy [8] et Kneser [6] qui a considéré des champs de 1-plans non forcément orientés sur le 2-tore ou la bouteille de Klein). Signalons aussi le travail remarquable de Seifert [9] qui a donné une classification complète des fibrations en cercles des variétés compactes de dimension 3, certaines fibres pouvant être exceptionnelles.

Les solutions des équations différentielles dans le domaine complexe sont des sous-variétés de dimension réelle 2 et peuvent être considérées comme les variétés intégrales d'un champ de 2-plans complètement intégrable. Voir à ce propos l'article de E. Ghys [82] où il relève que Painlevé [22] est l'un des premiers "à comprendre qu'il est préférable de considèrer les *solutions* comme des surfaces de Riemann non paramétrées qui sont en général non singulières mais dont les projections sur les axes de coordonnées présentent des singularités".

[3] Je remercie Daniel Bennequin et Alain Chenciner qui m'ont signalé ce fait qui est aussi mentionné dans le livre de B. L. Reinhart [71].

Dans un groupe de Lie G connexe, les translatés à droite d'une sous-algèbre de Lie de dimension p de l'algèbre de Lie de G (identifiée à l'espace tangent à G en l'élément neutre) est un champ de p-plans complètement intégrable ; les sous-variétés intégrales complètes sont les classes de G modulo un sous-groupe de Lie de G d'algèbre de Lie la sous-algèbre donnée.

Dans son livre *The theory of Lie groups* [13] paru en 1946, C. Chevalley appelle *distribution* ce que nous avons appelé plus haut *champ d'éléments de contact* et expose très soigneusement la caractérisation infinitésimale de l'intégrabilité (théorème dit de Frobenius) dans le cas analytique.

2 La thèse de Reeb

Nous allons suivre la genèse de la thèse de Reeb en examinant successivement les notes aux Comptes Rendus de l'Académie des Sciences publiées par Ehresmann et Reeb de 1944 à 1948.

Le problème de l'existence d'un champ de 2-plans complètement intégrable sur la 3-sphère avait été posé par Hopf en 1935 déjà sous la forme suivante (cf. [75], Tome I,1-2, p.507, et [70]) : existe-t-il sur la 3-sphère, munie de sa métrique riemannienne naturelle, un champ de vecteurs v partout non nul tel que le produit scalaire $v.rot(v)$ de v et de son rotationel s'annule ? Cette condition équivaut à la condition de complète intégrabilité mentionnée plus haut pour la forme de Pfaff correspondant à v par l'isomorphisme entre vecteurs et formes dans une variété riemannienne. Citons Reeb [70] : "Ce problème a aussitôt retenu l'attention de Ehresmann qui a mis ce problème "en conserve "dans un magnifique régistre-répertoire que mon maître m'avait autorisé à consulter en 1942, non sans apporter une certaine solennité -justifiée- à la communication du registre. Comment taire ici un regret : le riche contenu du répertoire ne semble pas avoir été porté à la connaissance d'un public élargi ?"

Ainsi pendant la guerre, à Clermont-Ferrand, où l'université de Strasbourg s'était repliée, Ehresmann propose à son élève Georges Reeb d'étudier les propriétés globales des champs d'éléments de contact complètement intégrables de dimension quelconque. On s'accorde à reconnaître que la note [11] soumise aux CRAS le 19 juin 1944 par Ehresmann et Reeb est l'acte de naissance de ce que Reeb appellera trois ans plus tard la théorie des variétés feuilletées.

Dans le premier paragraphe de cette note de deux pages, les auteurs rappellent la notion de champ K d'éléments de contact complètement intégrables sur une variété V, les variétés intégrales complètes étant définies comme les composantes connexes de V pour une topologie plus fine que celle donnée sur V (ayant comme base d'ouverts les variétés intégrales locales).

Dans le second paragraphe, ils montrent (théorème 1) que si V est la n-sphère ou l'espace euclidien de dimension n (et si p et $n-1$ sont pairs), la caractéristique d'Euler-Poincaré de toute variété intégrale compacte est nulle. C'est une observation importante (voir la fin de cette note), mais la démonstration qui en est donnée ici, et qui fait appel à des résultats de Gysin et Hopf, est curieusement indirecte. En effet Ehresmann avait déjà à sa disposition son théorème de relèvement des homotopies dans les espaces fibrés et il aurait suffit de remarquer que le fibré K restreint au complémentaire d'un point de S^n est trivial. Bien sûr

Ehresmann le remarquera très vite plus tard, et l'énoncé de ce théorème sera grandement généralisé, voir [18].

Dans le troisième paragraphe, les auteurs énoncent (Théorème 2) un résultat remarquable, que l'on désigne maintenant sous le nom de théorème de stabilité globale de Reeb en codimension 1, et dont voici l'énoncé.

Théorème 2 *Si K est un champ complètement intégrable d'éléments de dimension $n-1$ définissable par une forme de Pfaff*[4] *et si C_{n-1} est une variété intégrale compacte à groupe de Poincaré fini, toute variété intégrale suffisamment voisine est homéomorphe à C_{n-1} . Si de plus V_n est compact, toutes les variétés intégrales complètes sont homéomorphes à C_{n-1} et constituent un système de fibres de V_n. Si V_n est quelconque et si toutes les variétés intégrales complètes sont compactes, celles-ci sont homéomorphes et constituent un système de fibres de V_n.*

La démonstration fait intervenir des résultats de Bendixson sur les courbes définies par des équations différentielles."

Dans l'énoncé de ce théorème, il y a trois affirmations distinctes ; la première concerne les variétés intégrales complètes proches d'une variété intégrale compacte à groupe de Poincaré fini ; la seconde affirme que si la variété ambiante est compacte et s'il existe une variété intégrale compacte à groupe de Poincaré fini, alors toutes les variétés intégrales complètes sont les fibres d'une fibration de base un cercle (stabilité globale) ; enfin la troisième affirme que si toutes les variétés intégrales complètes sont compactes (pas forcément à groupe de Poincaré fini), alors elles sont les fibres d'une fibration (dont la base est un cercle ou une droite suivant que la variété ambiante est compacte ou non).

Notons que les deux premières affirmations n'ont aucun sens si les variétés intégrales sont de dimension 1 puisque dans ce cas une variété intégrale compacte est un cercle et a donc un groupe de Poincaré infini. Notons aussi que la brève indication sur la démonstration est très pertinente. Elle se réfère en effet aux résultats de Bendixson [4] sur les propriétés des familles de courbes définies par des équations différentielles dans le plan et beaucoup de ses raisonnements s'appliquent aux systèmes complètement intégrables en codimension 1. La structure d'ordre sur une petite courbe transverse aux variétés intégrales joue un rôle essentiel.

Enfin les auteurs décrivent le fameux exemple d'un champ de 2-plans complète-ment intégrable sur la 3-sphère (Ehresmann a précisé plus tard [28] que cet exemple était bien dû à Reeb), répondant ainsi par l'affirmative à la question de Hopf. Pour résoudre ce problème il fallait oublier la formulation analytique, et la reconsidérer d'une manière purement géométrique.

Voici une copie de ce dernier paragraphe.

" *Champs d'éléments plans ($p = 2$) complètement intégrables dans $\mathbb{S}^3$ ou $\mathbb{R}^3$.*

Toute surface intégrale compacte est homéomorphe au tore. Il existe effectivement des champs complètement intégrables admettant le tore comme surface intégrale. Par exemple dans le tore euclidien plein considérons la forme

$$rdr + (1 - r)d\phi$$

[4] Ceci équivaut à supposer que K est transversalement orientable.

r, distance au centre du cercle méridien correspondant ; ϕ, longitude ; θ, latitude. Cette forme est complètement intégrable et admet le tore $r = 1$ comme surface intégrale. En identifiant les bords de deux exemplaires de ce tore plein, les méridiens de l'un étant identifiés avec les parallèles de l'autre on obtient un champ complètement intégrable sur $\mathbb{S}^3$ admettant un tore comme surface intégrale. Ce tore est adhérent à toutes les autres surfaces intégrales complètes, et celles-ci sont toutes élémentaires et non compactes."

Ainsi les auteurs rappellent d'abord qu'une variété intégrale compacte doit être nécessairement un tore (Th. 1). Cette remarque est un bon point de départ pour la construction. On imagine le tore plongé de manière naturelle dans l'espace ; il borde de chaque côté un tore plein qu'il s'agit de remplir avec des surfaces, l'une d'entre elles étant le bord. Notons que la forme de Pfaff qui décrit le champ de 2-plans est seulement continue (analytique dans le tore plein si on remplace $1 - r$ par $1 - r^2$), mais pas de classe C^1 le long du bord commun des tores pleins. Néanmoins il existe un homéomorphisme de la 3-sphère sur elle-même transportant ces variétés intégrales sur les variétés intégrales d'un champ de 2-plans indéfiniment différentiable, mais pas analytique comme le remarque Reeb dans sa thèse [28] p.113 (c'est là qu'il pose le problème de l'existence d'un feuilletage analytique sur la 3-sphère). En effet si on coupe le tore avec un plan contenant un méridien, l'intersection de ce plan avec les variétés intégrales situées à l'extérieur sont des spirales s'enroulant asymptotiquement autour du méridien, alors que les intersections avec les variétés intégrales complètes situées à l'intérieur sont des cercles concentriques. Si l'on considère l'application de premier retour de Poincaré dans ce plan relativement à une petite transversale coupant le méridien, on trouve une application qui est l'identité sur la partie de la transversale qui est à l'intérieur du tore et qui n'est pas l'identité à l'extérieur ; une telle application ne peut être conjuguée à une application analytique.

Dans la note suivante de 1945 [12], Reeb montre qu'une variété intégrale compacte homologue à zéro a sa caractéristique d'Euler-Poincaré nulle. Il étudie également les variétés intégrales complètes d'une forme de Pfaff fermée partout non nulle sur une variété compacte en fonction de la dépendance sur les rationnels de ses périodes.

En 1946, Reeb publie une note [14] dans laquelle il considère le cas d'une forme de Pfaff ω complètement intégrable (c'est à dire telle que $\omega \wedge d\omega = 0$) ayant un zéro isolé à l'origine de $\mathbb{R}^n, n \geq 3$. Il remarque que si la matrice des dérivées partielles de cette forme à l'origine est définie (on dit que 0 est un centre), alors la forme admet un facteur intégrant. Il applique notamment le théorème de stabilité de la note de 1944 pour en conclure que si une forme intégrable sur une variété compacte de dimension $n \geq 3$ n'a comme points singuliers que des centres, cette variété est homéomorphe à une sphère et les variétés intégrales sont les variétés de niveau d'une fonction différentiable. (C'est un théorème beaucoup plus fort que celui que Milnor a utilisé dans son premier travail [32] sur les structures exotiques des sphères S^7, sous le nom de théorème de Reeb, et qui affirme qu'une variété compacte ayant une fonction de Morse avec seulement un minimum et un maximum est une sphère.)

En 1947 paraissent conjointement dans le même fascicule des Comptes Rendus de l'Académie des Sciences deux notes fondamentales, une d'Ehresmann *"Sur les espaces fibrés différentiables"* [16] et une de Reeb *"Variétés feuilletées, feuilles voisines"* [17].

Dans la première, Ehresmann définit la notion de fibré différentiable d'espace total E, base B, projection p de E sur B et fibre type F. Il énonce ensuite comme proposition

1 le fait fondamental qu'une submersion différentiable d'une variété compacte connexe E sur une variété connexe B est la projection d'une fibration. Il introduit ensuite la notion de connexion infinitésimale comme étant la donnée d'un champ d'éléments de contact de dimension égale à celle de la base et transverse aux fibres. Il remarque qu'un tel champ existe toujours (aussi dans le cas d'une submersion ; c'est en fait la clé de la démonstration de la proposition 1). Il montre ensuite que si ce champ est complètement intégrable et si les fibres de p sont compactes, alors les variétés intégrales complètes sont des revêtements de B et qu'on a une représentation du groupe de Poincaré de B basé en un point x dans un groupe d'automorphismes de la fibre au-dessus de x.

Dans la note qui suit, Reeb introduit pour la première fois la terminologie *feuille* (au lieu de *variété intégrale complète*). Il donne d'abord une définition (inspirée par Ehresmann, cf. [29], p.96) de la notion de structure de variété feuilletée de dimension p et codimension $q = n - p$ sur une variété topologique V de dimension n ; une telle structure est définie par un atlas formé de cartes qui sont des homéomorphismes d'ouverts de V sur des ouverts de l'espace numérique $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^q$, les changements de cartes appartenant au pseudogroupe des homémorphismes locaux de $\mathbb{R}^p \times \mathbb{R}^q$ localement de la forme

$$(x, y) \mapsto (g(x, y), h(y)), \quad x \in \mathbb{R}^p,\ y \in \mathbb{R}^q.$$

Les changements de cartes appliquent des morceaux ouverts de p-plans horizontaux homéomorphiquement sur des morceaux de p-plans horizontaux. Ils sont donc continus pour la topologie de $\mathbb{R}^n$ admettant ces morceaux comme base d'ouverts, d'où une topologie sur V qui en fait une variété de dimension p dont les composantes connexes sont les feuilles. L'ordre de différentiabilité des changements de cartes est par définition celui de la structure feuilletée. Cette définition oublie temporairement le champ d'éléments de contact et prend un sens précis dans le cas topologique. De plus elle permet une grande souplesse pour décrire des structures plus précises, en imposant aux changements de cartes d'appartenir à des sous-pseudogroupes particuliers, par exemple ceux qui pourraient préserver une structure complexe ou symplectique dans les variables y ou une structure riemannienne à courbure constante -1 dans les p-plans horizontaux, etc.

Une terminologie fort bien adaptée et suggestive a pris naissance[5]. Dans l'esprit de Reeb, localement une structure feuilletée est formée de sous-variétés parallèles comme les feuilles d'un livre ou encore mieux comme celles d'une pâte feuilletée (cf. Reeb [70]). Le terme "structure feuilletée" apparaît dans le Littré pour qualifier la structure de certaines roches. La traduction anglaise n'a pas été évidente. H. Samelson qui a référé pour les Math. Reviews tous les travaux de Reeb et Ehresmann jusqu'en 1955 traduit "feuille" par "leaf" et "variété feuilletée" par "leaved manifold" ou exceptionnellement "laminated manifold" (idem pour Chern qui a référé ma note de 1956). La traduction "foliation" a été utilisée par Ehresmann dans ses conférences données à Princeton[6] en 1953-54 (D.C. Spencer assistait à ces

[5] Dans son livre "Le Jargon des Sciences" (Paris, Hermann, 1966), l'écrivain Etiemble mentionne (p.76) plusieurs termes créés par les mathématiciens, dont "feuilleté". Je le cite : "Le pittoresque voulu de chacun de ces mots favorise l'intuition et seconde l'imagination." Et plus loin (p.77) : "*Théorie des variétés feuilletées, espace fibré différentiable*, vous êtes irréprochables.(...) Pour des concepts très neufs, on n'a pas eu recours au grec, ni à l'américain ; on s'est contenté d'ajouter un sens plus précis à tel vieux mot de la tribu ; on s'est contenté de faire comme Mallarmé, à d'autres fins."

[6] Voir le fac-similé reproduit au début de la partie II.2 de ses œ uvres complètes [75].

conférences et il est probable que cette traduction se soit dès lors imposée sur la côte Est des USA).

Dans cette note Reeb annonce aussi son fameux théorème de stabilité en codimension quelconque, à savoir que si une feuille F d'un feuilletage topologique est compacte et à groupe de Poincaré fini, alors toutes les feuilles rencontrant un voisinage convenable de F sont aussi compactes et à groupe de Poincaré fini. De plus si le feuilletage est deux fois différentiable, alors les feuilles voisines de F sont des revêtements de F.

Reeb s'est évidemment posé la question de savoir si le théorème de stabilité globale en codimension 1 avait un analogue en codimension supérieure à 1. Il y répond partiellement l'année suivante dans une note [20] en construisant un exemple de feuilletage de codimension 2 sur une variété compacte ayant des feuilles compactes simplement connexes et des feuilles compactes à groupe de Poincaré infini.

Enfin en 1948, il publie une note [21] où il démontre qu'une forme de Pfaff complète-ment intégrable holomorphe dans un voisinage de 0 dans $\mathbb{C}^n, n \geq 3$, et ayant un point singulier non dégénéré à l'origine admet une intégrale première holomorphe. Il en déduit le théorème analogue dans le cas analytique réel.

Reeb soutient sa thèse en 1948 et les détails des démonstrations seront publiés en 1952 dans un livre [29] contenant conjointement la thèse de Wu Wen-Tsun et la sienne.

3 Compléments dûs à Ehresmann

Autour de 1950, Ehresmann dégage une notion absolument essentielle pour les feuilletages qu'il appellera holonomie, et qui était sous-jacente, mais non explicitée, chez Reeb.

Dans son exposé au Colloque de Topologie de Bruxelles ([25], paragraphe 4), il reprend en les précisant les considérations de sa note de 1947 [16]. Il considère une connexion infinitésimale intégrable dans un fibré différentiable d'espace total E et de base connexe B, c'est-à-dire un champ d'éléments de contact transverse aux fibres de dimension égale à celle de la base. Il suppose que tout chemin dans la base peut se relever suivant un chemin horizontal pour la connexion, c'est-à-dire tangent aux éléments de contact du champ (c'est toujours le cas si la fibre F est compacte). Par relèvement des chemins, cette situation détermine un homomorphisme H du groupe fondamental de B basé en $x_0 \in B$ dans un groupe d'automorphismes de la fibre au-dessus de x_0, appelé l'*holonomie* de la connexion :

$$H : \pi_1(B, x_0) \to Aut(F).$$

Il remarque que la donnée de H permet de reconstruire à la fois le fibré E et le feuilletage transverse aux fibres, à savoir

$$E = \widetilde{B} \times F/(\tilde{x}, y) \text{ équivalent à } (\gamma.\tilde{x}, H(\gamma)(y)), \quad \gamma \in \pi_1(B, x_0)$$

où $\widetilde{B}$ est revêtement universel de B sur lequel le groupe fondamental opère par translations de revêtement. Le feuilletage transverse aux fibres est le quotient du feuilletage horizontal sur $\widetilde{B} \times F$.

Cette construction réciproque (appelée suspension de H) est fondamentale pour construire de multiples exemples de feuilletages intéressants et montre une relation intime entre actions par difféomorphismes de groupes discrets sur les variétés et théorie des feuilletages.

Il remarque que cette situation s'applique au cas où B est une feuille compacte d'un feuilletage différentiable admettant un système fondamental de voisinages tubulaires saturés par des feuilles ; les fibres d'un tel voisinage tubulaire assez petit seront transverses aux feuilles, et l'on aura donc un homomorphisme H du groupe fondamental de B en x dans le groupe des difféomorphismes d'un petit disque transverse aux feuilles centré en x et fixant x. Il note que, si toutes les feuilles voisines de B sont compactes, l'image de H est un groupe fini (grâce à un théorème de Montgomery) et conjugué à un sous-groupe du groupe orthogonal.

Ehresmann observe aussi dans cet article que le fibré tangent aux feuilles d'un feuilletage une fois différentiable a une connexion linéaire plate naturelle (il avait déjà fait cette remarque en 1948 dans [18]). Une connexion du fibré tangent à la variété feuilletée dont la restriction aux feuilles coïncide avec cette connexion plate est un des ingrédients essentiels dans la démonstration du théorème d'annulation de Bott [46]. L'autre est la théorie des classes caractéristiques et de l'homomorphisme de Chern-Weil admirablement expliqué par H. Cartan dans ce même colloque de topologie de Bruxelles [26] dont les deux exposés encadrent celui de Ehresmann. C'est dire que le théorème d'obstruction de Bott (1968) [46], qui se prouve en une demi-page à partir de ces deux ingrédients, aurait pu être démontré à cette époque déjà (cf. [81],volume 3, page XXVIII).

Dans une série de conférences données à l'Université de Princeton en 1953-54 dans le cadre du séminaire de Lefschetz, Ehresmann a généralisé comme suit ces considérations au cas d'une feuille quelconque B d'un feuilletage de classe C^r, où $r \geq 1$, de codimension q. Soit c un chemin dans B, c'est-à-dire une application continue de l'intervalle $[0, 1]$ dans B, et soient T et T' des petits disques de dimension q transverses aux feuilles centrés respectivement en $c(0) = x$ et $c(1) = x'$. On peut construire une famille continue de chemins c_y, paramétrée par les points y d'un petit voisinage U de x dans T, le chemin c_x étant égal à c et chaque chemin c_y étant contenu dans la feuille passant par y et reliant y à un point y' de T'.

L'application envoyant y sur y' est un difféomorphisme de classe C^r de U sur un voisinage U' de x' dans T'. Le fait fondamental à démontrer est que le germe (Ehresmann disait le jet local) de ce difféomorphisme en x ne dépend que de la classe d'homotopie de c dans B à extrêmités fixes. On en déduit donc en particulier, si $T = T'$, un homomorphisme H du groupe fondamental de B basé en x dans le groupe des germes en x des difféomorphismes locaux de T fixant x. Cet homomorphisme est appelé l'*holonomie* de la feuille B et son image est le *groupe d'holonomie* de B. Il est bien défini à conjugaison différentiable près si l'on change la section transverse T et le point base dans B. Il faut noter qu'à l'époque, la notion de germe n'était pas encore familière à tous !

Ehresmann remarque aussi que le théorème de stabilité de Reeb reste valable si l'hypotèse que la feuille compacte B a un groupe de Poincaré fini est remplacée par l'hypothèse que la feuille compacte a un groupe d'holonomie fini.

Plus généralement considérons une sous-variété T transverse aux feuilles, de dimension complémentaire, et coupant chaque feuille au moins une fois. On peut appliquer la construction précédente en faisant varier x et x' dans T et obtenir ainsi le pseudogroupe P d'holonomie du feuilletage associé à la transversale T. Le groupe d'holonomie de la feuille passant par le point x de T est conjugué au groupe des germes en x des éléments de P fixant x. Le pseudogroupe d'holonomie P' associé à une autre transversale T' vérifiant des propriétés

analogues est équivalent à P dans un sens naturel[7].

Après l'introduction par Ehresmann de l'holonomie, on peut dire que les fondements de la théorie des feuilletages étaient en place. Ehresmann n'a rédigé ses idées sur l'holonomie que trois ans plus tard dans une courte note aux CRAS publiée en collaboration avec Shih Weishu en 1956 [31] et dans un cadre beaucoup plus général (voir aussi [35], reproduit dans [75], tome II.2). C'est aussi dans cette note que le groupoïde d'holonomie d'un feuilletage (appelé parfois maintenant le graphe du feuilletage) est introduit pour la première fois, mais sans topologie (elle sera définie dans [35]).

4 De Reeb à Novikov

C'est en automne 1954 que je suis arrivé à Strasbourg comme boursier du gouvernement français pour étudier sous la direction d'Ehresmann. C'est René Thom qui m'a expliqué les idées d'Ehresmann sur l'holonomie durant l'hiver 54-55 (Ehresmann était souvent en voyage). J'ai réalisé assez vite que, dans le cas différentiable, l'holonomie d'une feuille propre F (c'est à dire telle que sa topologie de feuille coincide avec la topologie induite) permettait de reconstruire le germe de voisinage feuilleté de F et le caractérisait. L'année suivante j'ai démontré la non-existence de feuilletages analytiques de codimension 1 sur les variétés compactes simplement connexes [30], répondant en particulier par la négative à la question posée par Reeb concernant l'existence d'un feuilletage analytique sur la 3-sphère.

L'idée de la démonstration est simple. On observe tout d'abord qu'un feuilletage de codimension 1 sur une variété compacte V simplement connexe possède une feuille non compacte et qu'en général, en codimension 1, une feuille non fermée est toujours coupée par une courbe transversale fermée. Comme V est supposée simplement connexe, il existe une application différentiable f d'un disque D de dimension 2 dans V telle que la restriction de f au bord du disque soit la transversale fermée. On peut supposer f en position générale par rapport aux feuilles. On regarde alors la trace du feuilletage sur le disque : c'est un feuilletage de dimension 1 avec un nombre fini de singularités qui sont du type maxima-minima ou point selle. En utilisant les techniques classiques de Poincaré-Bendixson, on constate qu'il existe forcément une courbe fermée sur une feuille dont l'holonomie est l'identité d'un côté et une contraction de l'autre, ce qui contredit l'analyticité.

Après cela j'ai demandé à Ehresmann si les résultats que j'avais obtenus pourraient suffire pour une thèse. Il m'a répondu : " Oui si vous démontrez en plus que pour tout feuilletage de la 3-sphère, il existe une feuille compacte !". Malheureusement je n'ai pas réussi à répondre à cette question qui était dans l'air depuis longtemps, problème que Helmut Kneser avait abordé plusieurs fois sans succès et qui sera résolu par Novikov en 1964.

J'ai donc dû me contenter de rédiger en détails mes résultats partiels tout en essayant de comprendre plus profondément ce que pouvait signifier l'image réciproque sur un espace topologique K d'un feuilletage de codimension q et de classe C^r sur une variété V par une application continue f de K dans V (plus haut K était le disque unité D). Ceci m'a

[7] Cette relation d'équivalence évidente n'a été explicitée par écrit que beaucoup plus tard, lorsque cette notion s'est avérée nécessaire dans la formulation de certains énoncés (voir par exemple Bouma-Van Est, Indag. Math., 40 (1978), p. 313-347 et les papiers de Van Est qui ont suivi, ainsi que Haefliger, J. Diff. Geom. 15 (1980), p.269-347).

conduit à introduire la notion de Γ-structure sur K, où Γ était dans ce cas le groupoïde des germes de difféomorphismes de classe C^r de $\mathbb{R}^q$. Par exemple si f est l'inclusion d'une feuille F dans V, alors la structure induite est précisément l'holonomie de F. Ceci conduisait naturellement à mettre l'accent sur la structure transverse modelée par un pseudogroupe ou plus généralement un groupoïde topologique Γ . La notion de microfibré feuilleté associé à une Γ-structure apparaissait explicitement dans les deux cas particuliers traités en détails (holonomie d'une feuille et structure transversalement analytique de codimension un), mais les outils pour le cas général étaient présents. Malheureusement la rédaction [35] dans cette généralité est devenue parfaitement illisible. Pour cette raison, j'ai profité d'un cours au CIME à Urbino en 1962 [36] pour réexpliquer les résultats de la thèse de Reeb à la lumière des idées d'Ehresmann sur l'holonomie, et les résultats de ma thèse (en omettant la notion de Γ-structure, notion qui jouera douze ans plus tard [51] un rôle fondamental (voir les articles d'Anosov dans [84], volume 4, p.43-45 et 352-353).

Revenons en 1956. La thèse de Reeb était encore largement méconnue. A part un petit cercle autour d'Ehresmann, personne ne s'intéressait aux feuilletages. D'autres sujets passionnants émergeaient en topologie différentielle, par exemple la découverte par Milnor de structures différentiables exotiques sur des sphères ou l'étude des singularités des applications différentiables par Thom et Whitney. Dès 1958 mes intérêts se sont tournés vers la topologie différentielle et je ne suis revenu aux feuilletages de manière active qu'en 1969 quand j'ai eu connaissance de la remarquable découverte par Bott [46] d'une obstruction à la complète intégrabilité.

Jetons encore un bref coup d'oeil à la liste des publications concernant les feuilletages entre 1950 et 1963. A part celles d'Ehresmann et de ses élèves, notons le travail [27] très intéressant de Seifert (1950) qui montre qu'un feuilletage de dimension 1 sur $\mathbb{S}^3$ assez proche de la fibration de Hopf a toujours une feuille fermée et qui pose la question de savoir si un champ de vecteurs partout non nul sur $\mathbb{S}^3$ a toujours une trajectoire périodique (problème résolu par la négative beaucoup plus tard par P. Schweitzer en 1975 dans le cas C^1 et par Krystina Kuperberg en 1993 dans le cas C^∞, voir le rapport de E. Ghys [79]). Notons aussi le beau résultat de Milnor [34] en 1958 qui montre qu'un fibré vectoriel de rang 2 sur une surface orientable compacte S de genre ≥ 0 a une connexion linéaire plate, si et seulement si sa caractéristique d'Euler est bornée en valeur absolue par la moitié de la valeur absolue de la caractéristique d'Euler de S. En 1956 B. Reinhart soutient une thèse sous la direction de D. C. Spencer concernant les formes harmoniques sur les variétés feuilletées, et en 1959 il introduit la notion fondamentale de feuilletage riemannien [34], notion qui généralise en codimension supérieure celle de feuilletage donné par une 1-forme fermée étudiée dans la thèse de Reeb. Mentionnons aussi la thèse de R. Palais [33] en 1957 où la notion de feuilletage est utilisée dans l'étude des groupes de transformations (dans la lignée du livre de Chevalley [13]).

En 1963 Ehresmann et Reeb ont organisé un colloque sur les feuilletages à Grenoble qui a révélé un début d'intérêt outre-atlantique pour le sujet et plusieurs exposés ont été publiés dans les annales de l'Institut Fourier. C'est là que R. Sacksteder a publié ses premiéres contributions aux feuilletages, en particulier un exemple d'un ensemble minimal exceptionnel. Signalons aussi un travail remarquable de Sacksteder [39], (soumis en avril 1964), peut-être le premier où la notion globale de pseudogroupe d'holonomie (et non seulement de

groupe d'holonomie) est exploitée à fond. D'un théorème sur les pseudogroupes, il déduit en particulier que si un feuilletage de codimension 1 et de classe C^2 sur une variété compacte possède un minimal exceptionnel saturé par des feuilles, alors il existe un élément d'un groupe d'holonomie d'une feuille dont la dérivée est contractante.

La note de Novikov " Foliations of codimension 1 on manifolds " [37] parue en russe dans les Doklady en 1964, où il esquissait une démonstration que tout feuilletage de la 3-sphère par des surfaces possédait une feuille compacte a eu un impact considérable et a joué un rôle de catalyseur. Apparemment Novikov s'est intéressé aux feuilletages après la découverte par Anosov en 1962 des feuilletages naturels stables et instables associés au flot géodésique sur les variétés à courbure négative.

Dans cette note, Novikov commence par remarquer que, dans un feuilletage transversalement orienté de codimension 1 sur une variété compacte V, il existe une relation d'ordre partiel naturel dans l'ensemble des feuilles. La relation d'équivalence associée donne une décomposition de V en sous-ensembles saturés par des feuilles, qui peuvent être (théorème 1) soit réduits à une seule feuille compacte (si et seulement si cette feuille n'est pas coupée par une transversale fermée), soit un ouvert de V bordé par un nombre fini ≥ 0 de feuilles compactes (il se peut que cette composante soit V tout entier). Par exemple dans le feuilletage de Reeb, il y a 3 composantes : la feuille homéomorphe au tore et les intérieurs des deux tores pleins complémentaires, appelés depuis composantes de Reeb. Le théorème 2 affirme que si V est de dimension 3 et à groupe de Poincaré fini, alors il existe une feuille compacte qui borde une composante de Reeb. Si l'on cherche une telle composante, on doit pouvoir trouver dans V une courbe fermée C sur une feuille, non homotope à zéro sur cette feuille, qui peut être poussée d'un certain côté sur les feuilles voisines (c'est à dire l'holonomie de C est triviale de ce côté) suivant des courbes fermées homotopes à zéro sur ces feuilles. Une telle courbe sera appelée ici un cycle évanouissant. L'existence d'un tel cycle est le premier pas de la démonstration et n'est pas spécifique à la dimension 3 (sa démonstration pousse un peu plus loin l'argument du lemme fondamental de ma thèse). Une deuxième note [38], publiée quelques mois plus tard, généralise le théorème principal en affaiblissant les conditions pour l'existence d'un cycle évanouissant et précise qu'en dimension 3, si une feuille porte un cycle évanouissant, alors cette feuille est un tore bordant une composante de Reeb.

Les indications données par Novikov dans ces deux notes n'ont pas été suffisantes pour que nous puissions reconstruire la démonstration du théorème principal. Ce travail a pourtant eu tout de suite une grande influence. Par exemple, j'avais décrit en 64 les résultats de Novikov dans un colloque de topologie à Manchester et à la fin de l'exposé, Likorish avait déjà compris comment on pouvait feuilleter toutes les variétés de dimension 3 (voir [40]), résultat démontré indépendamment par Zieshang qui était en contact avec Novikov, voir [41]. Ce n'est que vers la fin de 1966 que j'ai pris connaissance du travail fondamental de Novikov (The Topology of Foliations [41]) qui contenait les détails des démonstrations (certains sont très subtils et ingénieux), et grâce à l'aide d'Israel Bernstein qui m'a traduit du russe des passages essentiels, j'ai pu exposer en détails la démonstration de l'existence d'une feuille compacte en janvier 67 dans le séminaire de " Current Literature " commun à l'Université de Princeton et l'Institute for Advanced Study. Les notes rédigées de cet exposé ont circulé à Berkeley notamment (et ont été traduites pour un séminaire Bourbaki [42] en 68). Novikov a fait une visite aux Etats-Unis en 1967, et lorsqu'il a passé à Berkeley il a été très surpris de rencontrer tant de jeunes mathématiciens travaillant sur les variétés

feuilletées. C'est dire qu'en quelques années, la situation avait radicalement changé et que la théorie des feuilletages devenait un sujet à la mode.

Simultanément en France, Harold Rosenberg avait commencé à étudier les feuilletages de codimension 1 sur les variétés de dimension 3 et les actions localement libres de $\mathbb{R}^n$ sur des variétés. Il avait obtenu de nombreux résultats, par exemple l'irréductibilité des 3-variétés feuilletées[8] sans composantes de Reeb [43] et commençait à diriger d'excellents élèves dans ces directions.

En relation directe avec la thèse de Reeb, signalons pour terminer la remarquable généralisation par Thurston [62] des théorèmes de stabilité (globale et locale) de Reeb, où l'hypothèse que le groupe de Poincaré d'une feuille compacte F est fini est remplacée dans le cas C^1 par l'hypothèse que $H^1(F, \mathbb{R}) = 0$ et que, en codimension 1 le feuilletage est transversalement orienté, et en codimension > 1 l'holonomie infinitésimale est triviale.

5 Quelques problèmes issus de la thèse de Reeb

Mentionnons quelques problèmes posés explicitement ou implicitement dans la thèse de Reeb.

1. Si une variété V admet un champ continu K d'éléments de contact de dimension p, admet-elle un feuilletage de dimension p ? (cf. Reeb [29], p.95). Plus précisément, K est-il homotope à un champ complètement intégrable ?

2. Existe-t-il un feuilletage analytique réel de codimension 1 sur $\mathbb{S}^3$? Tout feuilletage de codimension 1 de $\mathbb{S}^3$ possède-t-il toujours une feuille compacte ?

3. Dans un feuilletage topologique, les feuilles voisines d'une feuille compacte F simplement connexe sont-elles homéomorphes à F ? De manière équivalente, une submersion topologique propre est-elle une fibration ?

4. Si un feuilletage sur une variété compacte a toutes ses feuilles compactes, est-ce que chaque feuille possède un système fondamental de voisinages saturés par des feuilles, c'est-à-dire le groupe d'holonomie de chaque feuille est-il fini ? (ou de manière équivalente le volume des feuilles est-il borné ?)

5. A quelle condition une forme de Pfaff holomorphe complètement intégrable avec singularités admet-elle localement une intégrale première ?

Retraçons très brièvement l'évolution de ces problèmes. Comme nous venons de le voir, la première partie du deuxième problème a été résolue par Haefliger [30] en 1956 et la deuxième partie par Novikov [37] en 1964. Pour la solution du troisième problème, il a fallu attendre les résultats profonds de Kirby-Siebenmann sur l'extension des isotopies topologiques, voir Siebenmann [56]. Complétant des résultats de R. Moussu [68], Malgrange [69] a apporté une solution très satisfaisante au cinquième problème en montrant qu'une 1-forme holomorphe complètement intégrable admet un facteur intégrant au voisinage d'un point singulier si la codimension des points singuliers est au moins trois.

L'histoire du quatrième problème est très intéressante. D. Epstein a répondu d'abord par l'affirmative à la question 4 en 1972 dans le cas des variétés compactes de dimension 3 [57], en démontrant que, pour tout flot sans point fixe sur une variété de dimension 3

[8] Pour un survol des relations entre topologie des variétés de dimension 3 et feuilletages, voir Gabai [80].

ayant toutes ses orbites périodiques, il existe une action du cercle ayant les mêmes orbites. La démonstration de ce résultat est très délicate et ingénieuse, et les techniques initiées par David Epstein ont influencé tous les travaux ultérieurs sur ce probléme. Ainsi Edwards, Millett et Sullivan [66] ont montré que le résultat d'Epstein se généralisait en codimension 2, à savoir que, pour tout feuilletages de codimension 2 d'une variété compacte dont toutes les feuilles sont compactes, le volume des feuilles reste borné (ou de manière équivalente, tous les groupes d'holonomie sont finis). L'exemple par Sullivan [67] d'un flot sans point fixe sur une variété compacte de dimension 5 dont la longueur des orbites n'est pas borné a créé une grande surprise. Cet exemple a été suivi par beaucoup d'autres et finalement E. Vogt [73] a réussi le tour de force de construire un feuilletage en cercles de $\mathbb{R}^3$.

Le premier problème, celui de l'existence de feuilletages, est bien sûr fondamental et naturel. A l'époque de la thèse de Reeb on ne connaissait que peu d'exemples de feuilletages, et surtout de variétés compactes admettant des feuilletages de dimension > 1 (en dehors des fibrations). Comme on l'a vu plus haut, Lickorisch [40] et Zieschang [41] ont démontré indépendamment en 1965 par des constructions explicites l'existence de feuilletages de codimension 1 sur toute variété orientable de dimension 3 et J. Wood [47] a montré en 1969 que tout champ de 2-plans transversalement orientable sur une variété de dimension 3 est homotope à un champ complètement intégrable. Il a fallu attendre 1970 pour avoir les premiers exemples par Lawson [53] de feuilletages de codimension 1 sur des sphères de dimension impaire supérieure à trois, et l'année suivante (Tamura et Durfee) pour des exemples de feuilletages de codimension 1 sur toutes les sphères de dimension impaire (voir le survol de Lawson [61]). A. Phillips ([44] et [45]) a montré en 1968 que sur une variété connexe ouverte, tout champ de p-plans dont le complémentaire est un fibré trivial (ou plus généralement à groupe structural discret) était homotope à un champ complètement intégrable, en utilisant sa théorie des submersions.

La découverte par Raoul Bott [46] en 1968 de champs de p-plans non homotopes à un champ complètement intégrable a ouvert des perspectives insoupçonnées jusqu'alors. Bott montre que si $\mathcal{F}$ est un feuilletage de classe C^2 de dimension p et codimension q sur une variété M, alors les classes de cohomologie réelles de M qui sont des polynômes dans les classes de Pontrjagin du fibré vectoriel normal aux feuilles devaient s'annuler en degrés supérieurs à $2q$. Cette condition ne dépend que de la classe d'homotopie du champ de p-plans tangents aux feuilles. Comme on l'a mentionné plus haut, la démonstration de Bott est très directe et naturelle. A ma connaissance cette condition d'annulation est à ce jour la seule obstruction connue pour déformer un champ de p-plans en un champ complètement intégrable.

Cette découverte et les travaux de Phillips mentionnés ci-dessus m'ont motivé pour reprendre les idées de ma thèse et construire un espace classifiant $B\Gamma$ pour les Γ-structures (voir [50],[51] et [72], p. 82 pour un énoncé plus précis). En utilisant le théorème de transversalité de Gromov-Phillips [48] et [49], on y montre qu'il existe une théorie d'obstructions pour l'existence d'un feuilletage sur une variété connexe *ouverte* ayant une structure transverse modelée par Γ (supposé formé de germes de difféomorphismes locaux). Lorsque Γ est le groupoïde topologique Γ_q^r de tous les germes de difféormorphismes de $\mathbb{R}^q$ de classe C^r, on a une application naturelle $D : B\Gamma_q^r \to BGl_q$, induite par la différentielle, qu'on peut remplacer à homotopie près par une fibration. Le théorème d'annulation de Bott peut s'exprimer de la manière suivante : pour $r \geq 2$, l'homomorphisme $H^*(BGl_q, \mathbb{R}) \to H^*(B\Gamma_q^r, \mathbb{R})$ induit par

la projection D s'annule pour $* > 2q$. Thurston a démontré le résultat extrêmement remarquable suivant pour $r = \infty$: si K est un champ de p-plans de codimension q sur une variété M, il est homotope à un champ C^∞ complètement intégrable si et seulement si l'application de M dans BGL_q classifiant le fibré normal à K peut se relever suivant une application dans $B\Gamma_q^r$ (pour M compacte ou non et $q > 1$, voir Thurston [59] et aussi Eliashberg-Mischachev [83], et pour $q = 1$, voir Thurston [65]). Théoriquement on a donc en principe une théorie d'obstructions classique pour le problème $N1$ posé par Reeb. Tout la difficulté est de calculer le type d'homotopie de la fibre homotopique de D, notée $B\overline{\Gamma}_q^r$. Mather en 1971 déjà a découvert [52] que, pour $q = 1$, l'homologie de cette fibre était intimément liée à celle de l'homologie du groupe des difféomorphismes à support compact de $\mathbb{R}$ isotopes à l'identité (considéré comme groupe discret) ; ce résultat a été ensuite étendu par Mather et Thurston en toute codimension q. En particulier $B\overline{\Gamma}_q^r$ est $q+1$-connexe, si $r \neq q+1$ (voir le rapport de Mather [64] et ceux de Thurston [63] et [60]) ; c'est un résultat très profond qui équivaut à la simplicité du groupe des difféomorphismes à support compact de $\mathbb{R}^q$ isotopes à l'identité (dans [58] Thurston montre comment cette propriété peut être utilisée directement pour montrer qu'un champ de 2-plans est homotope à un champ complètement intégrable). Selon le même principe, Thurston a montré que le résultat de Mather [55] sur la trivialité de l'homologie du groupe des homéomorphismes à support compact de $\mathbb{R}^q$ entraîne qu'il n'y pas d'obstruction à déformer un champ de p-plans en un champ tangent à un feuilletage topologique. Entre-temps la découverte par Godbillon et Vey de leur fameux invariant [54] a révélé une connexion profonde entre la cohomologie de Gelfand-Fuks de l'algèbre de Lie des champs de vecteurs formels sur $\mathbb{R}^q$ et la cohomologie de la fibre de D et a permis de montrer la non-trivialité de plusieurs groupes de cohomologie de cette fibre. Malgré tous ces résultats spectaculaires et très profonds obtenus dans les années 70, le type d'homotopie de $B\overline{\Gamma}_q^r$ est encore très mal connu. Depuis cette époque, à part de rares exceptions (voir Tsuboi [76] et [77] qui a montré que $B\overline{\Gamma}_q^1$ est contractile), pratiquement aucun progrès n'a été réalisé. La solution des questions les plus naïves, par exemple la connexité de cette fibre pour $q > 1$ ou la détermination de son premier groupe d'homotopie non nul, semble hors de portée pour le moment.

On voit donc que toutes ces questions issues de la thèse de Reeb ont joué un rôle de catalyseur dans bon nombre de recherches sur les feuilletages ; la plupart d'entre elles ont été résolues dans les 30 ans qui ont suivi et Reeb, après avoir rappelé le scepticisme qui avait accueilli les débuts de la théorie des feuilletages, pouvait dire avec une fierté justifiée dans la préface du livre de Godbillon [78] : " Sur une durée de quarante années l'immeuble s'est édifié ; des centaines d'ouvriers ont oeuvré. L'édifice n'est pas achevé, mais on peut visiter. Oui, visiter est le mot".

Bibliographie

[1] G. Frobenius, *Ueber das Pfaffsche Problem*, J. Reine Angew. Math., vol **82** , 1877, p.230-315

[2] H. Poincaré, *Sur les courbes définies par des équations différentielles*, J. de Math., 1881-1882

[3] P. Painlevé, *Leçons sur la théorie analytique des équations différentielles professées à Stockholm*, 1895, in *Oeuvres* de Paul Painlevé, Tome 1, 1972

[4] I. Bendixson, *Sur les courbes définies par des équations différentielles*, Acta Matematica, vol **24**, 1901, p.1-80

[5] C. Carathéodory, *Untersuchungen über die Grundlagen der Thermodynamik*, Math. Ann., vol **67**, 1909, p.355-386

[6] H. Kneser, *Reguläre Kurvenscharen auf Ringflächen*, Math. Ann., vol **91**, 1924, p.135-154

[7] H. Hopf, *Vektorfelder in Mannigfaltigkeiten*, Math. Ann, vol 96, 1927, 225-250

[8] A. Denjoy, *Sur les courbes définies par des équations différentielles à la surface du tore*, J. de Math. Pures et Appl., vol **11**,1932,p.333-375

[9] H. Seifert, *Topologie dreidimensionaler gefaserter Räume*, Acta Matematica, vol **60**, 1932, 147-238

[10] W. Kaplan, *Regular curve families filling the plane*, Duke Journ., vol **7**,1940,p.153-185

[11] C. Ehresmann et G. Reeb, *Sur les champs d'éléments de contact complètement intégrables dans une variété continûment différentiable*, C. R. Acad. Sci. Paris, vol **218**, 1944, 955-957

[12] G.Reeb, *Sur les variétés intégrales des champs d'éléments de contact complètement inté-grables*, C. R. Acad. Sci. Paris, vol **220**, 1945, 236-237

[13] C. Chevalley, *Theory of Lie groups*, ed. Princeton University Press, 1946

[14] G.Reeb, *Sur les points singuliers d'une forme de Pfaff complètement intégrable*, C. R. Acad. Sci. Paris, vol **222**, 1946, 847-849

[15] C. Ehresmann, *Sur les sections d'un champ d'élements de contact dans une variété différentiable*, C. R. Acad. Sci. Paris, vol **224**, 1947, 444-445

[16] C. Ehresmann, *Sur les espaces fibrés différentiables*, C. R. Acad. Sci. Paris, vol **224**, 1947, 1611-1612

[17] G.Reeb, *Variétés feuilletées, feuilles voisines* , C. R. Acad. Sci. Paris, vol 224, 1947, 1613-1614

[18] C. Ehresmann, *Sur les variétés plongées dans une variété différentiable*, C. R. Acad. Sci. Paris, vol **226**, 1948, 1879-1880

[19] G.Reeb, *Sur les variétés feuilletées contenant une feuille compacte à groupe de Poincaré fini*, C. R. Acad. Sci. Paris, vol **226**, 1948, 1337-1339

[20] G.Reeb, *Sur les singularités d'une forme de Pfaff analytique complètement intégrable*, C. R. Acad. Sci. Paris, vol 227, 1948, 1201-1203

[21] G.Reeb, *Stabilité des feuilles compactes à groupe de Poincaré fini*, C. R. Acad. Sci. Paris, vol 228, 1949, 47-48

[22] H. Cartan, *Notions d'algèbres différentielles ; applications aux groupes de Lie et aux variétés où opère un groupe de Lie*, Colloque de Topologie, Bruxelles, 1950, 15-27

[23] C. Ehresmann, *Les connexions infinitésimales dans un espace fibré différentiable*, Colloque de Topologie, Bruxelles, 1950, 29-55

[24] H. Cartan, *La transgression dans un groupe de Lie et dans un espace fibré principal*, Colloque de Topologie, Bruxelles, 1950, 57-71

[25] H. Seifert, *Closed integral curves in* 3*-space and isotopic deformations*, Proc. Amer. Math. Soc., 1950, vol 1, 287-302

[26] C. Ehresmann, *Sur la théorie des variétés feuilletées*, Rend. Mat. e Appl., vol 10, 1951, 64-82

[27] G. Reeb, *Sur certaines propriétés topologiques des variétés feuilletées*, ed. Hermann, 1952

[28] A. Haefliger, *Sur les feuilletages analytiques*, C. R. Acad. Sc. Paris, vol 242, 1956, 2908-2910

[29] C. Ehresmann et S. Weishu, *Sur les espaces feuilletés, théorème de stabilité*, C. R. Acad. Sc. Paris, vol 243, 1956, 344-346

[30] J. Milnor, *On manifolds homeomorphic to the* 7*-sphere*, Ann. of Math., vol 64, (1956), 399-405

[31] R.S. Palais, *A global formulation of the Lie theory of transformations groups*, Memoirs Amer. Math. Soc., vol 22, 1957

[32] J. Milnor, *On the existence of a connection of curvature zero*, Comment. Math. Helv., vol 32, 1958, 215-223

[33] A. Haefliger, *Structures feuilletées et cohomologie à valeur dans un faisceau de groupoïdes*, Comm. Math. Helv., vol 32, 1958, 248-329

[34] B. L. Reinhart, *Foliated manifolds with bundle like metrics*, Ann. of Math., vol 69, 1959, 119-132

[35] C. Ehresmann, *Structures feuilletées*, Fifth Canadian Math. Congress, Montreal, 1961

[36] A. Haefliger, *Variétés feuilletées*, Ann. Mat. Scuola Norm. Sup. Pisa, vol 16, 1962, 367-397

[37] S. P. Novikov, *Foliations of codimension 1 on manifolds*, Dokl. Akad. Nauk SSSR, vol 155, 1964, 1010-1013

[38] S. P. Novikov, *Foliations of codimension 1 on manifolds*, Dokl. Akad. Nauk SSSR, vol 157, 1964, 788-790

[39] R. Sacksteder, *Foliations and pseudogroups*, Amer. J. Math., vol 87, 1965, 79-102

[40] W. B. R. Lickorish, *A foliation for* 3*-manifolds*, Ann. of Math., vol 81, 1965, 414-420

[41] S. P. Novikov, *The topology of foliations*, Trudy Moskov. Mat. Obshch, vol 14, 1965, 248-278

[42] A. Haefliger, *Travaux de Novikov sur les feuilletages*, Séminaire Bourbaki, 1967-68, 1968

[43] H. Rosenberg, *Foliations by planes*, Topology, vol 7, 1968, 131-138

[44] A. Phillips, *Foliations on open manifolds I*, Comm. Math. Helv., vol 43, 1968, 204-211

[45] A. Phillips, *Foliations on open manifolds II*, Comm. Math. Helv., vol 44, 1969, 367-370

[46] R. Bott, *On a topological obstruction to integrability in Global Analysis*, Berkeley 1968, ed. Proc. Sympos. Pure Math. Vol XVI, 1970, 127-131

[47] J.W. Wood, *Foliations on* 3*-manifolds*, Ann. of Math., vol 89, 1969, 336-358

[48] M. L. Gromov, *Stable mappings of foliations into manifolds*, Izv. Akad. Nauk SSSR, vol 33, 1969, 707-734

[49] A. Phillips, *Smooth maps transverse to a foliation*, Bull. Amer. Math. Soc., vol 76, 1970, 792-797

[50] A. Haefliger, *Feuilletages sur les variétés ouvertes*, Topology, vol 9, 1970, 183-194

[51] A. Haefliger, *Homotopy and integrability*, in *Manifolds*, Amsterdam, 1970, Springer Lecture Notes in Math., 1971, vol 197, 133-163

[52] J. Mather, *On Haefliger's classifying space*, Bull. Amer. Math. Soc., vol 77, 1971, 1111-1115

[53] H.B. Lawson, *Codimension one foliations of spheres*, Ann. of Math., vol 94, 1972, 494-503

[54] C. Godbillon et J. Vey, *Un invariant des feuilletages de codimension* 1, C. R. Acad. Sc., Paris, vol 273, 1971, 92-95

[55] J.N. Mather, *The vanishing of the homology of certain groups of homeomorphisms*, Topology, vol 10, 1971, 297-298

[56] L.C. Siebenmann, *Deformation of homeomorphisms on stratified sets I, II*, Comm. Math. Helv., 1972, vol 47, 123-163

[57] D. B. A. Epstein, *Periodic flows on three manifolds*, Ann. of Math, vol 95, 1972, 66-82

[58] W.P. Thurston, *A local construction of foliations for three-manifolds*, in *Differential Topology*, Stanford, 1973, Proc. Sympos. Pure Math., vol 27-1, 1975, 315-319

[59] W.P. Thurston, *The theory of foliations in dimension greater than one*, Comm. Math. Helv., vol 49, 1974, 2314-231

[60] W.P. Thurston, *Foliations and group of diffeomorphisms*, Bull. Amer. Math. Soc., vol 80, 1974, 304-307

[61] H.T. Lawson, *Foliations* , Bull. Amer. Math. Soc. , vol 80, 1974, 369-418

[62] W.T. Thurston, *A generalization of the Reeb stability theorem*, Topology, vol 13, 1974, 347-352

[63] W.P. Thurston, *On the construction and classification of foliations*, Proceedings of the ICM, Vancouver, 1974, tome 1, 547-549

[64] J.N. Mather, *Foliations and local homology of groups of diffeomorphisms*, in *Proceedings of the ICM*, Vancouver, 1974, tome 2, 35-37

[65] W.P. Thurston, *Existence of codimension one foliation*, Ann. of Math., vol 104, 1976, 249-268

[66] R. Edwards, K.C. Millet, D. Sullivan, *Foliations with all leaves compact*, Topology, vol 16, 1977, 13-32

[67] D. Sullivan, *A counterexample to the periodic orbit conjecture*, Publ. Math. Inst. Hautes Etudes Sci., vol 46, 1976, 5-14

[68] R. Moussu, *Sur l'existence d'intégrales premières pour un germe de forme de Pfaff*, Ann. Inst. Fourier, vol 26-2, 1976, 171-220

[69] B. Malgrange, *Frobenius avec singularités, I. Codimension 1*, Publications Mathématiques I.H.E.S., vol 46, 1976, 163-173

[70] G. Reeb, *Structures feuilletées*, in *Differential Topology, Foliations and Gelfand-Fuks cohomology*, Rio de Janeiro, 1976, Springer Lecture Notes in Math., vol 652, 1978, 104-113

[71] B. L. Reinhart, *Differential Geometry of Foliations*, in*Ergebnisse der Mathematik und ihrer Grenzgebiete*, vol 99, Publ. Springer Verlag

[72] A. Haefliger, *Holonomie et classifiants*, in *Structure transverse des feuilletages*, Toulouse 1982, Astérisque, vol 116, 1984, 70-97

[73] E. Vogt, *A foliation of* $\mathbb{R}^3$ *and other punctured 3-manifolds by circles*, Publ. Math. IHES, vol 69, 1989, 215-232

[74] E. Ghys, *L'invariant de Godbillon-Vey*, in *Séminaire Bourbaki*, Vol. 1988/89, Astérisque, vol 177-178, 1989, 155-181

[75] C. Ehresmann, *Oeuvres complètes et commentées, I-1,2 Topologie algébrique, topologie diffé-rentielle. II.2 Catégories ordonnées. Applications des ordres en topologie*, ed. Andrée Charles Ehresmann, Cahiers de Topologie Géom. Différentielle, Suppléments aux volumes 23 et 24., 1984, 1982

[76] T. Tsuboi, *On the connectivity of the classifying spaces for foliations*, in *Algebraic Topology* (Evanston, II, 1988), Contemp. Math., vol 96, 1989, 319-331 ed. Amer. math. Soc., Publ. Providence, RI

[77] T. Tsuboi, *On the foliated products of class* C^1, Ann. of Math., vol 130, 1989, 227-271

[78] C. Godbillon, *Feuilletages, Études géométriques*, ed. Birkäuser, 1991

[79] E. Ghys, *Constructions de champs de vecteurs sans orbite périodique (d'après Kristyna Kuperberg)*, in *Séminaire Bourbaki*, Vol. 1993/94, Astérisque, vol 227, 1995, 283-307

[80] D. Gabai, *Foliations and* 3*-manifolds*, Proc. ICM Kyoto 1990, Publ. Math. Soc. Japan, 1991, 609-619

[81] R. Bott, *Collected Papers*, vol 3, ed. Robert D. MacPhearson, 1994-95, Publ Birkhäuser

[82] E. Ghys, *Les systèmes dynamiques holomorphes*, in*Dynamique et géométrie complexe* (Lyon 1997), Panor. et Synthèses, vol 8, Publ Soc. Math. France, Paris, 1999, 1-10

[83] Y. Eliashberg and N.M. Mishachev, *Wrinkling of smooth mappings III, Foliations of codimension greater than one, Topological Methods Nonlinear Analysis*, vol 11, 1998, 321-350

[84] In Encyclopedia of Mathematics. An updated and annotated translation of the Soviet. Mathematics Encyclopedia, Publ Kluwer, Academic Publishers

La redécouverte des spineurs par les mathématiciens dans la seconde moitié du XXème siècle

Jean Pierre Bourguignon
(CNRS-IHÉS et École polytechnique) IHES
35, route de Chartres
F-91440 Bures-sur-Yvette

and

(École polytechnique)
Centre de Mathématiques
UMR 7640 du CNRS
F-91128 PALAISEAU
(France)

1 Les spineurs avant 1930

La première apparition des spineurs est due en 1913 au systématisme d'Élie Cartan qui, en procédant à la classification des représentations complexes de l'algèbre de Lie so_n, remarque l'existence de représentations exotiques, i.e. ne provenant pas des représentations dans les espaces de tenseurs construits sur $\mathbf{R}^n$. Dans [21], il passe au cas réel et remarque que ces représentations sont bivaluées pour le groupe SO_n, indiquant qu'il doit exister un groupe qui revêt de façon double le groupe des rotations dont ces représentations exotiques sont de vraies représentations. C'est en 1928 qu'insatisfait de l'utilisation de l'équation de Klein-Gordon comme substitut à l'équation de Schrödinger pour les électrons, Paul-Adrien-Maurice Dirac propose (cf. [24]) une nouvelle équation invariante sous l'action du groupe de Lorentz $\mathrm{SO}_{3,1}$ qui régirait le mouvement des électrons relativistes. Pour ce faire il procède par coefficients indéterminés et cherche un opérateur différentiel du premier ordre $\mathcal{D}$ (qui depuis s'appelle l'*opérateur de Dirac"*) sous la forme $\mathcal{D} = \sum_{i=0}^{3} A_i\, \partial/\partial x^i$, où les A_i sont

des matrices à coefficients complexes agissant sur la fonction d'ondes vectorielle (c'est une innovation considérable) représentant l'électron et les x^i sont les coordonnées locales de l'espace-temps, x^0 représentant le temps, et les x^i $(1 \leq i \leq 3)$ les coordonnées d'espace. Comme $\mathcal{D}^2$ doit redonner l'opérateur des ondes classique, encore appelé *d'Alembertien*, $\mathcal{D}$ doit nécessairement vérifier

$$\mathcal{D}^2 = -\partial^2/(\partial x^0)^2 + \partial^2/(\partial x^1)^2 + \partial^2/(\partial x^2)^2 + \partial^2/(\partial x^3)^2 \ ,$$

ce qui donne, entre les coefficients de l'opérateur différentiel, les fameuses relations de commutation

$$A_i\, A_j + A_j\, A_i = \pm\, 2\, \delta_{ij} \ , \ 0 \leq i, j \leq 3 \ ,$$

où δ_{ij} désigne le *symbole de Kronecker généralisé* qui vaut -1 lorsque $i = j = 0$, $+1$ lorsque $i = j > 0$, et 0 lorsque $i \neq j$. Ces relations sont celles que vérifient les générateurs de l'algèbre de Clifford associée au produit scalaire de Minkowski sur $\mathbf{R}^4$ pris comme modèle de l'espace-temps. Cette algèbre s'identifie à une algèbre de matrices opérant sur un *autre* espace de dimension 4 (sur le corps des nombres complexes celui-là), qui est précisément celui des *spineurs*[1].

Il est à noter que dès 1927, Wolfgang Pauli plaidait pour que l'on attribue un moment magnétique propre à l'électron. Comme l'équation relativiste proposée par Dirac a donné des résultats en accord remarquable avec l'expérience, cette approche a immédiatement eu du succès du côté des physiciens. Il y avait pourtant de considérables problèmes à surmonter car le fait que l'opérateur soit du premier ordre lui faisait avoir des valeurs propres négatives ce qui, physiquement, signifiait l'existence de solutions d'énergie négative, la marque théorique de l'anti-matière. L'existence de ce nouveau type de matière allait être établie expérimentalement quelques années plus tard.

La révolution fondamentale qui s'opère alors place durablement les vecteurs de ces espaces de représentation exotiques au cœur de la physique moderne : la fonction d'ondes de l'électron ne peut pas être une fonction scalaire mais un champ de spineurs, puisque c'est le nom qui est donné à ces objets. (On peut relier ce nom au moment magnétique propre qui les rend sensibles au choix d'un axe de rotation orienté, ce qui est cohérent avec le fait de prendre un revêtement à deux feuillets du groupe des rotations.)

On peut dire que c'est alors que s'opère une cassure entre mathématiciens et physiciens : les spineurs ont trouvé une place fondamentale en physique alors qu'en mathématiques ils restent (pour quelque temps encore !) des objets marginaux.

[1] Les physiciens, eux, utilisent plutôt la notation γ pour désigner ces matrices et pour le cas de dimension 4 ont souvent recours à leur expression explicite en termes de *matrices de Pauli*. Le point de vue consistant à faire apparaître les spineurs comme espace de représentation des algèbres de Clifford est aussi celui qui est maintenant systématiquement utilisé par les mathématiciens pour présenter les spineurs. La jonction avec les espaces de représentation du groupe Spin vient naturellement du fait que ce groupe se présente dans ce cadre comme un sous-groupe du groupe multiplicatif de l'algèbre de Clifford (cf. [23] ou [33] pour une présentation systématique).

2 Des années 30 aux années 50 chez les mathématiciens

Pendant cette période, on continue à noter des articles épisodiques sur les spineurs par des mathématiciens, écrits soit par des algébristes purs, soit par des mathématiciens qui s'intéressent à la physique mathématique. C'est pourtant au cours de cette période de maturation que vont se développer tous les outils indispensables pour le grand retour des spineurs sur le devant de la scène mathématique. Passons-les rapidement en revue.

La généralisation de la notion de spineur à toutes les dimensions est faite de façon systématique par Richard Brauer et Hermann Weyl dans l'article [20] paru en 1935. L'espace des spineurs complexes attachés à un espace complexe de dimension n est de dimension $2^{\left[\frac{n}{2}\right]}$. On peut noter en passant que la dimension 4 est la seule où l'espace des spineurs est de même dimension que l'espace sur lesquels on les construit. La classification des algèbres de Clifford (sur le corps des nombres complexes) fait apparaître une périodicité d'ordre 2 : cette algèbre est une algèbre simple lorsque la dimension est paire, et une somme de deux algèbres simples lorsque la dimension est impaire. Un fait algébrique élémentaire joue alors un rôle considérable : toute algèbre simple est une algèbre de matrices sur un espace vectoriel. Dans notre cas, c'est précisément là qu'apparaissent les spineurs.

En 1937 Élie Cartan publie sa monographie très complète sur la théorie des spineurs [22], mais on doit noter que les développements qu'il y donne sont presque complètement algébriques. Peut-être plus important encore : le livre se termine sur l'énoncé d'un *théorème* qui, s'il est mal compris, peut même donner l'impression qu'une *géométrie des spineurs* dans un espace riemannien général est impossible. Élie Cartan dit précisément :

Theorem 2.1 – *Avec le sens géométrique que nous avons donné au mot "spineur", il est impossible d'introduire des champs de spineurs selon la technique riemannienne classique : en effet, ayant choisi un système de coordonnées locales x^i pour l'espace, on ne peut pas représenter un spineur par un nombre fini N de composantes u_α, aussi grand soit-il, de telle sorte que sa dérivation covariante s'écrive*

$$u_{\alpha;i} = \frac{\partial u_\alpha}{\partial x^i} + \Lambda_{\alpha\ i}^{\beta}\, u_\beta \ ,$$

où les $\Lambda_{\alpha\ i}^{\beta}$ sont des fonctions bien définies des x^h.

En 1954 paraît une autre monographie qui a aussi donné un fondement algébrique indiscutable à la théorie des spineurs : c'est le texte de Claude Chevalley *"Algebraic Theory of Spinors"* récemment republié (cf. [23]).

Parallèlement de nouveaux points de vue, souvent implicites chez Élie Cartan , apparaissent en géométrie, qui vont constituer le cadre de pensée de toute la géométrie de la deuxième moitié du XXème siècle : il s'agit de la théorie des *fibrés* présentée notamment dans le livre de Norman Steenrod [46] et les articles de Charles Ehresmann (cf. [25]).

Les travaux de Friedrich Hirzebruch sur le théorème de Riemann-Roch, et notamment son livre [30], jettent les bases d'un grand rapprochement entre des domaines des mathématiques qui étaient restés jusque là très éloignés. Cette fusion ne pourra prendre toute son ampleur qu'en prenant des lunettes spinorielles pour la lire.

3 Le tournant du théorème de l'indice

3.1 L'énoncé

Autour de 1960 les travaux de Michael F. Atiyah et Isadore M. Singer sur le théorème de l'indice vont fournir le cadre à un rapprochement spectaculaire de plusieurs branches des mathématiques, notamment géométrie et analyse, le mot "géométrie" devant être pris ici aussi bien dans son acception la plus classique incluant la géométrie différentielle, spécialement dans ses développements riemanniens, que le point de vue défendu par Henri Poincaré la rapprochant de la *topologie*.

Dans la série d'articles [7], [8], [10], ils prouvent (mais la première preuve publiée se trouve dans [39]) un théorème qui va avoir des conséquences importantes dans un grand nombre de sujets, y compris, sur la fin du siècle, en physique théorique. Le voici.

Theorem 3.1 *de l'indice d'Atiyah-Singer– Si, sur une variété différentielle compacte M, P est un opérateur différentiel linéaire envoyant les sections d'un fibré vectoriel E sur les sections d'un fibré vectoriel F qui est supposé elliptique*[2]*, alors*

$$\text{indice}_{\text{anal}} P = \text{indice}_{\text{géom}} P$$

où $\text{indice}_{\text{anal}} P = \dim \ker P - \dim \operatorname{coker} P$ *et où* $\text{indice}_{\text{géom}} P = \int_M f(\operatorname{Symb} P)$ *pour une fonction bien définie f du symbole principal de P noté ici Symb P.*

Des extensions à des situations où un groupe de Lie agit et au cas de familles d'opérateurs elliptiques ont été considérées très tôt par Atiyah et Singer dans [9] et [11] avec de nombreuses applications géométriques, topologiques, et même algébriques.

Deux des arguments essentiels de la preuve tournent autour de l'invariance par homotopie : d'une part le fait que l'indice géométrique puisse recevoir une interprétation cohomologique établit l'invariance par homotopie de ce membre de l'égalité et d'autre part la possibilité donnée par le calcul pseudo-différentiel (cf. par exemple [44]) d'établir l'*invariance par homotopie de l'indice analytique des opérateurs elliptiques* garantit que les phénomènes fondamentaux du point de vue de l'indice peuvent être lus dans le symbole principal. Ceci incite à chercher dans chaque classe d'homotopie un opérateur particulier pour lequel l'établissement de la preuve serait plus simple. Cette approche ne prendra toute sa force qu'un peu plus tard avec les travaux utilisant une approche plus fonctorielle et universelle (cf. [3] inspiré par Isadore Singer [45] et Peter Gilkey [26]) et par Vijay K. Patodi dans le cas holomorphe (cf. [41]).

3.2 L'apparition de l'opérateur de Dirac et des spineurs

C'est à l'occasion de cette recherche qu' Atiyah et Singer redécouvrent l'opérateur fondamental agissant sur les champs de spineurs qu'est l'opérateur de Dirac dans le contexte

[2] On rappelle qu'un opérateur différentiel entre sections de deux fibrés vectoriels E et F sur une variété différentielle est dit *elliptique* si, en tout point x de M, ses termes de plus haut degré de différentiation donnent lieu, pour toute direction cotangente, à une application linéaire bijective entre les fibres E_x et F_x en ce point.

riemannien. Cette "redécouverte" est suffisamment spectaculaire pour que Blaine Lawson et Marie-Louise Michelsohn dans leur monographie fondamentale sur ce sujet [33] donne le nom d'*opérateur d'Atiyah-Singer* à l'opérateur de Dirac dans ce cadre[3]. Comment la jonction se fait-elle ? Tout simplement parce qu'il apparaît possible d'engendrer toutes les classes d'homotopie de symboles d'opérateurs elliptiques à partir de celui de l'opérateur de Dirac en tordant le fibré des spineurs par des fibrés auxiliaires. Établir cette propriété est bien entendu une question purement topologique. D'où l'importance stratégique prise par la preuve du théorème de l'indice pour les opérateurs de Dirac.

Rappelons tout d'abord que la possibilité sur une variété différentielle M de définir globalement des champs de spineurs (on dit alors que la variété est *spin*) est conditionnée par l'annulation de la seconde classe de Stiefel-Whitney de M, $w_2(M) \in H^2(M, \mathbf{Z}/2\mathbf{Z})$. Par ailleurs, si, sur M supposée orientable et spin, on se donne une métrique riemannienne g, il est alors possible de définir la notion de *structure spinorielle* sur M, supposée de dimension n, comme le choix d'une classe d'équivalence de fibrés principaux à groupe structural Spin_n dont un représentant revêt le fibré SO_n-principal des repères g-orthonormés. Grâce à cette construction et à l'interprétation de la dérivation covariante comme forme de connexion à valeurs dans l'algèbre de Lie du groupe structural, il est possible de contourner l'objection que semblait faire Cartan : on ne représente pas les champs de spineurs dans des bases naturelles associées à des systèmes de coordonnées quelconques mais dans des repères mobiles qui respectent la métrique riemannienne. La définition générale de l'opérateur de Dirac $\mathcal{D}$ agissant sur un champ de spineurs ψ est donnée par la formule

$$\mathcal{D}\psi = \sum_{i=1}^{n} e_i . D_{e_i}\psi$$

où D désigne la dérivation covariante déduite de la connexion de Levi-Civita associée à la métrique g et . la multiplication de Clifford (celle par laquelle il est possible de faire agir les formes sur les spineurs, vus comme modules sur l'algèbre de Clifford). Dans la direction cotangente ξ au point x, le symbole principal de $\mathcal{D}$ agissant sur un spineur ψ au point x vaut

$$\mathrm{Symb}\,(\mathcal{D})(\xi)(\psi) = \xi.\psi \ ,$$

ce qui assure que $\mathcal{D}$ est bien un opérateur elliptique. (En effet le carré de son symbole principal dans la direction cotangente ξ est l'application $\psi \mapsto \xi.\xi..\psi = -g(\xi,\xi)\,\psi$ à cause de l'identité fondamentale de la multiplication dans les algèbres de Clifford, et cette application est bien un isomorphisme lorsque g est riemannienne et $\xi \neq 0$.)

3.3 Premières applications

Parmi les premières applications du théorème de l'indice figurent des théorèmes d'intégralité d'un certain nombre d'invariants topologiques. En effet un calcul direct permet d'établir que $\mathrm{indice}_{\mathrm{géom}}\,\mathcal{D} = \hat{A}(M)$, où le $\hat{A}$-*genre* est une combinaison des classes caractéristiques de Pontrjagin donnée par une des suites multiplicatives de Hirzebruch (cf. [30]). En basses dimensions le calcul du $\hat{A}$-genre de M supposée *spin* donne de façon automatique des théorèmes

[3]Sur ce point on lira avec intérêt la préface de cet ouvrage.

de divisibilité remarquables : ainsi, à cause de la formule

$$\hat{A}(M^4) = -\frac{1}{24}\, p_1(M^4) \left(= \frac{1}{8}\, \text{signature}(M^4) \right) \ ,$$

on obtient que *la signature d'une variété spin compacte de dimension 4 est divisible par 8* (qui est une version affaiblie en dimension 4 du théorème de Rokhlin[4]) et la formule

$$\hat{A}(M^8) = \frac{1}{2^7\, 3^2\, 5} \left(-4\, p_2(M^8) + 7\, p_1^2(M^8) \right) \ ,$$

où p_i désigne la ième classe de Pontrjaguine, donne d'autres propriétés de divisibilité (en fait les dénominateurs croissent très vite comme on le pressent en voyant ce qui se passe en passant de la dimension 4 à la dimension 8). Il est ainsi possible de disposer d'une vaste généralisation des résultats obtenus par Friedrich Hirzebruch dans le cadre holomorphe à propos du théorème de Riemann-Roch.

4 Les à-côtés topologiques

A partir et autour du théorème de l'indice va se développer dans toute sa généralité la *K-théorie*, une théorie cohomologique fondée sur les relations d'équivalence que l'on peut bâtir entre fibrés vectoriels sur une variété. Cela a conduit à raffiner le $\hat{A}$-genre, et à calculer l'anneau de spin-cobordisme (deux variétés munies de structures spinorielles sont *spin-cobordantes* si leurs structures spinorielles peuvent être induites à partir de la structure spin d'une variété réalisant un cobordisme entre elles).

Il y a plusieurs façons d'aborder la K-thorie, et le livre [1] est une bonne introduction. Nous prendrons ici plutôt comme fil directeur l'article [4] qui relie une construction initialement due à Alexander Grothendieck associant des groupes aux modules de Clifford réels à la KO-théorie des sphères (i.e. la K-théorie des fibrés vectoriels réels, le groupe orthogonal portant toute la topologie du groupe linéaire réel puisqu'il en est un rétracte), les modules de Clifford complexes étant eux reliés à la K-théorie des sphères. (On pourra aussi consulter [32].) Il en ressort une nouvelle preuve du théorème de périodicité de Bott pour les groupes orthogonaux et unitaires. Le calcul de l'anneau de spin-cobordisme se révèlera un peu plus tard contenir des informations géométriques insoupçonnées, qui pourront être utilisés dans des contextes inattendus.

Dans ce contexte le *caractère de Chern ch* associé à un fibré vectoriel E sur une variété M est une classe de cohomologie à coefficients rationnels sur M obtenue à partir de la classe totale de Chern $c(E) = 1 + c_1(E) + ... + c_n(E) = \prod_{k=1}^{n}(1 + c_k)$ par la construction

$$ch(E) = \mathrm{e}^{x_1} + ... + \mathrm{e}^{x_n} \ .$$

(Remarquer que pour un fibré en droites complexes E, $ch(E) = \mathrm{e}^{c_1(E)}$.) Il joue ici un rôle important, ainsi que les *suites multiplicatives* de Hirzebruch comme la suite de Todd, la

[4]Dans la section 5 nous donnerons une façon de retrouver toute la force du théorème de Rokhlin en utilisant plus finement l'information contenue dans les spineurs harmoniques, i.e. ceux qui sont annihilés par l'opérateur de Dirac.

suite de la classe $\hat{A}$ ou la suite de Hirzebruch (cf. [30]) proprement dite. Ces combinaisons de classes caractéristiques ont des interprétations qu'il est possible de relier avec le théorème de l'indice. Voici une forme possible de ces relations qui donne un nouveau contexte pour comprendre à un certain niveau de généralité le Théorème de l'Indice.

Theorem 4.1 *(Version cohomologique du Théorème de l'Indice, cf. [9])– Soit P un opérateur elliptique linéaire opérant entre fibrés vectoriels sur une variété compacte M orientée de dimension n et soit* $\mathrm{Symb}(P) \in K_0(TX)$ *la classe du symbole de P (ici K_0 désigne la K-théorie à support compact). Alors*

$$\mathrm{indice}_{\mathrm{anal}} P = (-1)^{\frac{n(n+1)}{2}} \{\pi!(ch(\mathrm{Symb}P))\, \hat{\mathbf{A}}(M)^2\}[M]$$

où $\pi : TM \longrightarrow M$ est la projection naturelle et $\hat{\mathbf{A}}$ désigne la $\hat{A}$-classe totale.

Nous avons déjà présenté certains théorèmes d'intégralité qui auparavant appartenaient complètement au domaine de la topologie et qui, par l'éclairage nouveau donné par le théorème de l'indice, prennent une nouvelle dimension plus géométrique.

D'autres questions *a priori* topologiques prennent aussi un autre éclairage comme les questions d'obstruction au plongement dans des espaces modèles (cf. [33] page 281) et la recherche du nombre maximum de champs de vecteurs linéairement indépendants en chaque point d'une variété particulière (cf. [2]). Ces questions ont trouvé des prolongements récents dans la théorie des genres elliptiques, elle aussi un temps à la mode chez les physiciens théoriciens (cf. [52].)

Pour chacune de ces questions des constructions inspirées des algèbres de Clifford liées à la théorie des fibrés, et notamment à leur lecture "spinorielle", donnent les résultats optimaux, ce qui prouve la force de pénétration exceptionnelle du point de vue spinoriel dans ces différents domaines des mathématiques.

5 Les à-côtés géométriques

Avant de rentrer dans des choses plus inattendues, on peut déjà mentionner une remarque de nature géométrique qui permet de montrer que le théorème de divisibilité de la signature d'une variété spin de dimension 4 peut être améliorée d'un facteur 2 (la signature est nécessairement divisible par 16) : par le théorème de l'indice, la dimension de l'espace des spineurs harmoniques donne la valeur du $\hat{A}$-genre ; or, en dimension 4, cet espace porte naturellement une structure symplectique, donc est nécessairement de dimension paire, d'où le gain d'un facteur 2.

5.1 Spineurs harmoniques et existence de métriques à courbure scalaire positive

De façon assez inattendue, le théorème de l'indice pour l'opérateur de Dirac peut être utilisé pour obtenir des énoncés globaux de géométrie riemannienne, dans le cadre des variétés spin bien entendu[5].

[5]En fait il est possible d'affaiblir cette condition et de considérer le cas où $w_2(M)$ n'est pas nulle mais la restriction modulo 2 d'une classe entière. Pour cela il faut considérer la notion plus générale de structure

La formule fondamentale qui permet de faire le lien entre la courbure d'une variété riemannienne supposée spinorielle et les champs de spineurs est due à Erwin Schrödinger (cf. [43]). Elle a été redécouverte par André Lichnerowicz (cf. [34]) dès l'apparition du théorème de l'indice. Elle s'énonce ainsi : Formule de Schrödinger-Lichnerowicz.– *Pour tout champ de spineurs ψ sur une variété spinorielle*

$$\mathcal{D}^2\psi = D^*D\psi + \frac{1}{4}\,\mathrm{Scal}_g\,\psi\ .$$

où $Scal_g$ désigne la courbure scalaire de la métrique riemannienne g et D^* l'adjoint de l'opérateur de dérivation covariante[6].

La formule de Schrödinger-Lichnerowicz relie les spineurs harmoniques à la courbure scalaire puisqu'après intégration sur la variété M et intégration par parties elle assure que, si $\mathcal{D}\psi = 0$, $\int_M(||D\psi||^2 + \frac{1}{4}\,\mathrm{Scal}_g\,||\psi||^2) = 0$. Lorsque ψ n'est pas identiquement nul, une telle égalité n'est pas possible si $\mathrm{Scal}_g > 0$ en tout point de M . Comme l'existence de spineurs harmoniques non triviaux est forcée par la non-nullité du $\hat{A}$-genre de M, on déduit de ces considérations le théorème suivant.

Theorem 5.1 *(cf. [34])– Toute variété spinorielle compacte dont le $\hat{A}$-genre n'est pas nul n'admet pas de métrique à courbure scalaire positive.*

Cette jonction a été considérablement approfondie dans les travaux de Mikhael Gromov et H. Blaine Lawson qui l'ont prolongée dans deux directions très importantes :

- ils ont ainsi montré dans [27] que dans le cas des variétés simplement connexes la question de l'existence d'une métrique à courbure positive ne dépendait que de la classe de spin-cobordisme (avec la conséquence que si une variété n'est pas spinorielle, alors elle admet toujours une telle métrique) grâce à des chirurgies appropriées en codimension 2 ;
- dans le cas des variétés qui ont au contraire un très grand groupe fondamental permettant de prendre des revêtements d'ordre arbitrairement élevé (comme c'est le cas pour le tore par exemple), en appliquant une variante de la formule de Schrödinger-Lichnerowicz à des sections ramenées d'un fibré spinoriel sur une sphère par une application bien choisie, ils montrent dans [28] que ces variétés n'admettent pas de métrique à courbure scalaire positive.

Gromov et Lawson ont étendu leur approche au cas de variétés non compactes faisant la connexion avec une conjecture de Serguei Novikov (cf. [29]). Les développements autour de cette question ont pris une extension considérable, mettant en évidence des liens profonds entre la topologie algébrique des variétés et l'existence de métriques ayant une courbure positive. Les résultats les plus complets ont été obtenus par Stefan Stolz (cf. [47], [48], et [49]).

Spin_c qui permet de tordre le fibré des spineurs par un fibré en droites complexes. Noter que c'est toujours le cas sur une variété complexe car alors on a toujours $w_2(M) = c_1(M) \bmod 2$, où $c_1(M)$ désigne la première classe de Chern de M. On peut alors définir globalement un opérateur de Dirac associé à la structure Spin_c.

[6] Noter que cette formule est un calcul formel dans lequel la signature de la métrique n'entre pas, ce qui assure que ce calcul est aussi valable pour des métriques lorentziennes, cas qui intéressait bien entendu Schrödinger.

5.2 Une autre approche du théorème de la masse positive en Relativité Générale

L'utilisation de l'opérateur de Dirac, notamment via la formule de Schrödinger-Lichnerowicz, s'est encore révélée très fructueuse à l'interface entre la géométrie différentielle et la physique théorique.

C'est ainsi qu'Edward Witten a proposé (cf. [51]) une preuve de la *conjecture de la masse positive* en Relativité Générale utilisant un spineur harmonique ayant un comportement à l'infini prescrit (pour une preuve plus complète, on peut consulter [40]). Cette conjecture consiste à établir, dans le contexte d'un espace-temps muni d'un champ de gravitation asymptotiquement plat, la positivité d'un invariant géométrique assez subtil mesurant, pour un système relativiste isolé, la *masse totale* du système comme intégrale de certains dérivées du champ métrique sur des sphères de plus en plus grandes dans une hypersurface du type espace (où la métrique est donc riemannienne). Les hypothèses physiques concernent la positivité de la densité d'énergie qui s'exprime par des inégalités ponctuelles sur les composantes spatiales et temporelles du tenseur d'impulsion-énergie (i.e. le membre de droite des équations d'Einstein de la Relativité Générale).
Cette approche se limite bien entendu aux variétés spinorielles, mais l'autre approche qui utilise les sous-variétés minimales est, elle, limitée à cause des problèmes de régularité des solutions de cette équation en dimension supérieure ou égale à 8.

5.3 L'approche spinorielle de l'espace-temps de Roger Penrose et les twisteurs

Il n'est pas possible de quitter ce domaine géométrique sans citer les efforts considérables faits par Roger Penrose pour repenser l'ensemble de la physique, et notamment la gravitation, en partant des *spineurs*. C'est le but poursuivi dans les deux volumes [42].

Pour lui une des leçons fondamentales de la Mécanique Quantique est de mettre les structures complexes à la base de la formulation de la physique. C'est notamment pourquoi il a proposé de remplacer l'espace-temps à 4 dimensions par l'espace de toutes les structures complexes sur l'espace-temps, son *espace de twisteurs*. En dimension 4, l'espace des structures complexes s'identifie à la droite projective complexe formée sur l'espace (de dimension 2 complexe) des spineurs complexes par la correspondance algébrique qui à un spineur ψ associe la structure complexe J définie par $JX.\psi = i\,X.\psi$, où . désigne la multiplication de Clifford et i désigne l'imaginaire fondamental. Sur une variété il convient de prendre l'espace fibré dont la fibre en chaque point est l'espace des twisteurs de l'espace tangent en ce point-là.

Pour décider si l'espace des twisteurs porte lui-même une structure complexe, par exemple en résolvant les équations locales permettant de trouver un système de coordonnées complexes, on se trouve confronté à un système sur-déterminé, qui est associé à un opérateur différentiel P, appelé l'*opérateur des twisteurs*, qui transforme un champ de spineurs ψ en un champ de 1-formes-spineurs par la formule $P_X\psi = D_X\phi + \frac{1}{n}X.\mathcal{D}\psi$, où D désigne la dérivation covariante de Levi-Civita, et X un vecteur tangent en un point de la variété de base M (la notation $\mathcal{D}$ pour l'opérateur a déjà été introduite auparavant).

Parmi les sous-produits de cette approche on doit signaler la linéarisation de plusieurs

problèmes non-linéaires particulièrement importants pour la physique des théories de jauge sur des espaces standards : c'est ainsi que la recherche de solutions auto-duales des équations deYang-Mills sur la sphère S^4 est traduit par cette construction en la recherche de sections holomorphes sur le fibré remonté sur l'espace projectif $P\mathbf{C}^3$, qui est l'espace des twisteurs de S^4. D'autres situations ayant une grande symétrie peuvent aussi être traitées en ces termes (cf. [5] et [6]).

5.4 L'approche par l'équation de la chaleur et le théorème de l'indice des familles

Petit à petit, il s'est révélé que la géométrie différentielle permettait de pénétrer encore plus profond dans la compréhension du Théorème de l'Indice en en fournissant une version locale, où la représentation de la classe de cohomologie qui porte l'indice par une forme différentielle bien particulière était une information précieuse. L'idée fondamentale est d'*utiliser le développement asymptotique de l'équation de la chaleur*, et de donner des arguments a priori pour que les termes plus singuliers que celui qui porte de l'information sur l'indice puissent être négligés dans le calcul. La formule fondamentale est la suivante.

Theorem 5.2 – *Soit P un opérateur différentiel elliptique envoyant linéairement les sections d'un fibré E dans celles d'un fibré F sur une variété riemannienne compacte M. Alors*

$$\mathrm{ind}_{\mathrm{anal}} P = \mathrm{trace}(\mathrm{e}^{-tP^*P}) - \mathrm{trace}(\mathrm{e}^{-tPP^*})$$

pour tout $t > 0$, où P^* désigne l'adjoint de P construit en utilisant des produits scalaires sur E et F.

Cette approche était déjà en germe dans les travaux de Herbert McKean et Isadore M. Singer [45], et ont été conduits jusqu'au bout par Ezra Getzler (cf. [15]) et Jean-Michel Bismut (cf. [16], [17]) parmi d'autres.

L'intérêt des physiciens théoriciens pour ces questions a engendré une énorme littérature sur cet aspect du sujet dont il est difficile de rendre compte ici tellement elle est volumineuse.

6 La supersymétrie et la supergéométrie

Les physiciens théoriciens ont franchi une nouvelle étape dans leur utilisation des champs de spineurs avec l'introduction du concept, dont la pertinence pour la physique elle-même est encore une spéculation. Pour expliquer que les familles de particules élémentaires que l'on a pu identifier dans la nature sont aussi peu nombreuses, il s'agit d'affirmer qu'il existe une symétrie supplémentaire dans la nature, une *supersymétrie*. Ce nouveau type de symétries transforment des particules représentées par des champs de spineurs en particules représentées par des champs de tenseurs classiques.

Ce point de vue amène à rediscuter toutes les symétries de la physique, et bien entendu à introduire un supergroupe de Poincaré pour discuter des symétries de l'espace-temps minkowskien. La question de l'existence de supersymétries sur des espaces-temps courbes est plus subtile car elle conduit à considérer des variétés qui admettent des champs de *spineurs*

ψ, dits *de Killing*, solutions du système surdéterminé $D_X\psi + \frac{1}{n}\lambda X.\psi = 0$. Ce système exprime que ψ est à la fois une solution de l'équation des twisteurs mentionnée dans la section précédente et un spineur propre pour l'opérateur de Dirac pour la valeur propre λ. Le caractère surdéterminé apparaît par le fait que la métrique doit être automatiquement d'Einstein, et doit vérifier d'autres identités. Le problème de classification des variétés compactes admettant des spineurs de Killing a été complètement résolu par Christian Bär (cf. [13]), après des travaux préliminaires de Peter Van Nieuwenhuizen et N.P. Warner (cf. [50]).

Ceci a rejailli et a permis de découvrir qu'un certain nombre de structures géométriques avaient une expression particulièrement simple et efficace lorsqu'elles étaient traduites dans le langage des spineurs. On pourra notamment consulter à ce propos les travaux [37], et [38] d'Andrei Moroianu.

Dans la mouvance des réflexions des physiciens, ces approches ont suggéré de développer une *supergéométrie*, qui aurait, à côté du secteur traditionnel où les variables de position commutent, un secteur *fermionique* où les variables de position anticommutent. Dans ce contexte on peut définir des *super-variétés*, des *super-champs de vecteurs*, des *super-métriques*, etc. Il convient donc de reformuler tous les outils traditionnels de la géométrie dans ce contexte élargi. C'est l'école russe qui a poussé cette approche le plus loin (cf. par exemple [14], et [35] mais il existe maintenant une littérature immense sur ce point de vue). Une des généralisations intéressantes, introduite originellement par Daniel Quillen et bien exploitée par Jean-Michel Bismut, est celle d'une *superconnexion* (cf. [18]).

Il est même possible d'introduire un aspect arithmétique à ces questions en faisant jouer un rôle particulier aux anneaux d'entiers. Le point crucial est de considérer toutes les variables sur un pied d'égalité. Ceci a conduit à un nouveau point de vue dans l'étude des équations diophantiennes, qui s'intègre bien avec le point de vue dit d'Arakelov (cf. [36]). On rejoint là une question qui a été un peu laissé dans l'ombre, à savoir que de nombreuses variantes du Théorème de l'Indice font appel à des éléments de torsion dans les groupes où ils sont définis, qui ne sont donc pas calculables par le biais de formes différentielles. Pour rendre compte de ces propriétés il est nécessaire de faire appel à d'autres techniques, certaines plus algébriques.

Bibliographie

[1] M.F. Atiyah*K-Theory*, With notes by Anderson, D.W., Adv, Book Classics, Addison-Wesley. **xviii**, Redwood City, CA, 1989.

[2] M.F.Atiyah *Vector Fields on Manifolds,* Arbeitsgemeinschaft für Forschung des Landes Nordrhein-Westfalen **200** (1970), 7–23.

[3] M.F. Atiyah, R. Bott, V.K. Patodi*On the Heat Equation and the Index Theorem*, Inventiones Math. **19** (1973), 279–330 ; Erratum, ibidem **28** (1975), 277–280.

[4] M.F. Atiyah, R. Bott, A. Shapiro,*Clifford Modules*, Topology **3** Suppl. 1 (1964), 3–38.

[5] M.F. Atiyah, N. Hitchin, I.M. Singer,*Self-duality in four-dimensional Riemannian Geometry*, Proc. Royal Soc. London **362** (1978), 425–461.

[6] M.F. Atiyah, N. Hitchin, Y. Manin, V. Drinfeld,*Construction of Instantons*, Phys. Lett. **65A** (1978), 185–187.

[7] M.F. Atiyah, I.M. Singer,*The Index of Elliptic Operators on Compact Manifolds*, Bull. Amer. Math. Soc. **69** (1963), 422–433.

[8] M.F. Atiyah, I.M. Singer, *The Index of Elliptic Operators, I*, Ann. Math. **87** (1968), 484–530.

[9] M.F. Atiyah, I.M. Singer*The Index of Elliptic Operators, III*, Ann. Math. **87** (1968), 546–604.

[10] M.F. Atiyah, I.M. Singer,*Index Theory for Skew-Adjoint Fredholm Operators*, Publ. Math. Inst. Hautes Études Sci. **37** (1969), 5–26.

[11] M.F. Atiyah, I.M. Singer,*The Index of Elliptic Operators, IV*, Ann. Math. **93** (1971), 119–138.

[12] M.F. Atiyah, I.M. Singer,*The Index of Elliptic Operators, V*, Ann. Math. **93** (1971), 139–149.

[13] C. Bär,*Real Killing Spinors and Holonomy,* Commun. Math. Phys. **154** (1993), 509–521.

[14] F.A. Berezin, *Introduction to Algebra and Calculus with Anticommuting Variables*, Moscow University Press, Moscow, 1983 (in Russian).

[15] N. Berline, E. Getzler, M. Vergne,*Heat Kernels and Dirac Operators,* Grundlehren Math. Wiss. **298**, Springer-Verlag, Berlin-Heidelberg-New York, 1992.

[16] J.M. Bismut, *The Atiyah-Singer Theorems : A Probabilistic Approach I*, J. Functional Anal. **57** (1984), 56–99 ; *idem, II,* ibidem, 329–348.

[17] J.M. Bismut, *The Atiyah-Singer Index Theorem for Families of Dirac Operators : Two Heat Equation Proofs*, Inventiones Math. **83** (1986), 91–151.

[18] J.M. Bismut, *Superconnexions, indice local des familles, déterminant de la cohomologie et métriques de Quillen*, Mém. Soc. Math. France **46** (1991), 27–72.

[19] J.P. Bourguignon, P. Gauduchon,*Spineurs, opérateurs de Dirac et variations de métriques,* Commun. Math. Phys. **144** (1992), 581–599.

[20] R. Brauer, H. Weyl,*Spinors in n dimensions*, Amer. J. Math. **57** (1935), 425–449.

[21] É. Cartan,*Les groupes projectifs qui ne laissent invariante aucune multiplicité plane,*. Bull. Soc. Math. France **41** (1913), 53–97.

[22] É. Cartan,*Leçon sur la théorie des spineurs,* Gauthier-Villars, Paris, 1937 ; $2^{\text{ème}}$ édition, *The Theory of Spinors*, Hermann, Paris, 1966.

[23] C. Chevalley,*The Algebraic Theory of Spinors*, Columbia Univ. Press **VIII**, New York, 1954 ; nouvelle édition in *Collected Works*, Vol. **2**, *The Algebraic Theory of Spinors and Clifford Algebras*, Cartier, P. et Chevalley, C. (éd.). Springer, Berlin-Heidelberg, 1997.

[24] P.A.M. Dirac,*The Quantum Theory of the Electron*, Proc. Roy. Soc. **A 117** (1928), 610.

[25] C. Ehresmann, *Sur la théorie des espaces fibrés*, in *Topologie albébrique*, Colloque Intern. Centre Nat. Rech. Sci. **12** (1949), 3–15.

[26] P. Gilkey, *Curvature and the Eigenvalues of the Laplacian for Elliptic Complexes* , Adv. Math. **10** (1973), 344–382.

[27] M. Gromov, H.B. Lawson, *The Classification of Simply Connected Manifolds of Positive Scalar Curvature*, Ann. Math. **111** (1980), 423–434.

[28] M. Gromov, H.B. Lawson, *Spin and Scalar Curvature in the Presence of a Fundamental Group. I.*, Ann. Math. **111** (1980), 209–230.

[29] M. Gromov, H.B. Lawson, *Positive Scalar Curvature and the Dirac Operator on Complete Riemannian Manifolds.*, Publ. Math. Inst. Hautes Études Sci. **58** (1983), 295–408.

[30] F. Hirzebruch, *Neue Topologische Methoden in der algebraischen Geometrie*, Ergebnisse Math. Grenzgebiete **9**, Springer-Verlag, Berlin-GöttingenHeidelberg, 1956.

[31] L.P. Hughston, R. Penrose, P. Sommers, M. Walker, *On a Quadratic First Integral for the Charged Particle Orbits in the Charged Kerr Solution*, Commun. Math. Phys. **27** (1972), 303–308.

[32] M. Karoubi, *Algèbres de Clifford et K-théorie,* Ann. Sci. Éc. Norm. Sup. Paris **1** (1968), 161–270.

[33] H.B. Lawson, M.L. Michelsohn,*Spin Geometry*, Princeton Math. Series **38**, Princeton Univ. Press, Princeton, 1989.

[34] A. Lichnerowicz,*Spineurs harmoniques*, C.R. Acad. Sci. Paris **257** (1963), 7–9.

[35] Yu.I. Manin. *Gauge Fields and Complex Geometry*, Nauka, Moscow, 1984.

[36] Yu.I. Manin, *New Dimensions in Geometry*, Lect. Notes in Math. **1111**, Springer-Verlag, 59–101.

[37] A. Moroianu, *Géométrie spinorielle et groupes d'holomomie*, Cours Peccot, Collège de France (1998).

[38] A. Moroianu, *$Spin_c$-manifolds and complex contact structures*, Commun. Math. Phys. **193** (1998), 661–673.

[39] R.S. Palais, *Seminar on the Atiyah-Singer index theorem*, Ann. of Math. Studies **57**, Princeton Univ. Press, Princeton, 1965.

[40] T. Parker, C.H. Taubes, *On Witten's Proof of the Positive Energy Theorem*, Commun. Math. Phys. **84**, (1982), 223–238.

[41] V.K. Patodi, *Curvature and the Fundamental Solution of the Heat Operator*, J. Indian Math. Soc. **34** (1970), 269–285.

[42] R. Penrose, W. Rindler, *Spinors and Space-Time*, Vol I and II, Cambridge Univ. Press, Cambridge, 1986.

[43] E. Schrödinger, 1932.

[44] R. Seeley, *Complex Powers of an Elliptic Operator*, Amer. Math. Soc. Proc. Symp. Pure Math. **10** (1967), 288–307.

[45] I.M. Singer, *Future Extensions of Index Theory and Elliptic Operators,* in *Prospects in Mathematics*, Proc. Sympos. Princeton, Ann. Math. Studies **70**, Princeton Univ. Press, Princeton, 1971, 171–185.

[46] N. Steenrod, *The Topology of Fibre Bundles* **VIII**, Princeton Math. Series, Princeton Univ. Press, Princeton, 1951.

[47] S. Stolz, *Simply Connected Manifolds of Positive Scalar Curvature*, Bull. Amer. Math. Soc. **23** (1990), 427–432.

[48] S. Stolz, *Simply Connected Manifolds of Positive Scalar Curvature*, Ann. Math. **136** (1992), 511–540.

[49] S. Stolz, *Positive Scalar Curvature Metrics. Existence and Classification Questions*, Proc. Intern. Congress Math. 94, Chatterji, S.D. (ed.), Birkhäuser, Basel, 1995, 625–636.

[50] P. Van Nieuwenhuizen, N.P. Warner, *Integrability Conditions for Killing Spinors*, Commun. Math. Phys. **93** (1984), 277–284.

[51] E. Witten, *A new Proof of the Positive Energy Theorem*, Commun. Math. Phys. **80**, (1981), 381–402.

[52] E. Witten, *The Index of the Dirac Operator in Loop Space*, in *Elliptic Curves and Modular Forms in Algebraic Topology*, Princeton 1988, Lect. Notes in Math. **1326**, Springer-Verlag, Berlin, 161–181.

Local spinor structures in V. Fock's and H. Weyl's work on the Dirac equation (1929)

Erhard Scholz
Wuppertal

1 Introduction

In the early 20th century the most important impact of mathematical physics on geometry came from relativity theory. Historical and philosophical questions of this interplay have been discussed at various occasions.[1] The rise of quantum physics brought about a second, philosophically, technically and conceptually much deeper shift for the relationship of geometry to physics. It started in the late 1920s, gained momentum in the second half of the past century and started to dominate the image of knowledge for the deeper levels of physical geometry during its last two decades.[2] Other contributions to these conference proceedings are evidence for the actuality of this recent and ongoing shift in our understanding of physical geometry, which is far from completed and continues to be an open-ended and controversial project.[3]

An important turn in the relationship between relativity, quantum mechanics and field theory, which also sheds light on the nature and role of geometry in this conceptual complex, has been initiated by Hermann Weyl and Vladimir Fock in early 1929. They both started to investigate (generalized) Dirac fields in the context of general relativity by the introduction of local spinor structures on Lorentz manifolds. This topic has been taken up anew in the 1960s from a global point of view ; the role of the Dirac operator for the interplay between differential geometry and topology in the last third of the century is being discussed in J. P. Bourguignon's contribution to this volume.

Up to the end of the 1920s mathematical physicists had essentially two symbolic tools for the represention of physical fields at their disposal : vectors/tensors (including differential

[1] Among them [1] [20].
[2] For a first historical exploration see [3], in particular J. Stachel's introductory remarks.
[3] Cf. contribution of M. Atiyah to this volume.

forms) and linear connections (mostly but not always affine), most important among them, of course, the Levi-Civita connection of general relativity (GRT). After 1918 H. Weyl tried to convince physicists and mathematicians for some time to use another type of connection (*length connection*) in combination with a conformal (class of) Lorentz metric in his first, strictly metrical gauge geometry. Most physicists who considered Weyl's length connection at all referred to it as just another differential 1-form $\varphi = \sum \varphi_i dx^i$ with a peculiar, perhaps even strange, transformation behaviour.

In the early 1920s A. S. Eddington started to build his attempts towards a unified field theory of electromagnetism, gravitation and matter using general affine connections (not necessarily derived from a metric); and Einstein joined him for a while from 1923 onward. These activities were part of a broader move towards unified field theories (UFT's) with a first high tide in the 20s of the last century, which has been studied historically, among others, by Vladimir Vizygin [39] and, more recently and in a different methodological approach, by Catherine Goldstein and Jim Ritter [17].[4] V. Vizgin presents the relationship of UFT and quantum physics (QP) as one of competing research programs mutually influencing each other. The introduction of local spinor structures by Fock and Weyl in 1929 is a beautiful example for his case. Both, Weyl and Fock, were struck by the early successes of the Dirac equation for the explanation of the motion of the electron and attempted an integration of GRT and the Dirac field. In such an attempt they were not alone. Other authors, like Wiener and Vallarta, attempted a similar integration along different lines, building upon Einstein's recent theory of "distant parallelism". They attempted to adapt the Dirac field to a framework of classical UFT's that soon turned out to be too restrictive.

Weyl and Fock, the latter after an initial phase of sympathizing with distant parallelism, pursued an approach of a covariant differentiation of spinor fields derived from the underlying Levi-Civita connection, in contrast to the distant parallelism program. Both realized that, in doing so, an underdetermination of the ensuing spinor connection led naturally to an additional $U(1)$-symmetry. They used the latter for a representation of the electromagnetic field comparable to, although slightly different from, Weyl's earlier approach using a length connection. Thus they arrived at a geometric-analytical structure in which the actual knowledge of gravitation, electromagnetism and the basics of the quantum theory of the moving electron could be represented in an integrated form.[5] Both authors posed the question how geometry might be brought into agreement with quantum physical knowledge of their time. They arrived at strongly diverging evaluations as to what they had achieved in this respect and what geometrization of quantum physics might mean at all (last section).

Before I discuss Weyl's and Fock's respective approaches and differences with respect to "quantum geometry", I want to sketch the background of common knowledge from which they started and outline their 1929 work.

[4] Another high tide, in a different historical/scientific context and with changed conceptual/symbolical approaches, started in the 1970s. It has not yet found the detailed and critical historical investigation it deserves, although work has started [2], [28], [16], [29].

[5] For a discussion of Weyl's 1929 work on gravitation and the electron see also [37].

2 Setting the stage in the later 1920's for Weyl and Fock

During the 1920s the constitutive conditions for the mathematization of geometry and matter changed deeply. In the middle of the decade (1925/26) the "new" quantum mechanics took shape, with its different, in central aspects compatible, although at least historically and conceptually not completely equivalent, versions put forward by Heisenberg/Born/Pauli, Schrödinger and Dirac.[6] In immediate continuation of this turn, in late 1926, W. Heisenberg started to investigate the symmetry of atomic electrons using surprisingly oldfashioned mathematics, Serret's *Algèbre supérieure* from 1879. But already in the following year the two young Hungarians E. Wigner and J. von Neumann, working in Berlin and Göttingen, applied group representation methods for this goal, as did H. Weyl in a lecture course devoted to this subject in the winter semester 1927/28 at the ETH Zürich[7].
Still in 1926, W. Pauli attempted to characterize the new hypothetical electron "spin" in terms of quantum mechanical symbolism and introduced a pair of "wave" functions $(\psi_1(x), \psi_2(x))$, $x \in \mathbb{R}^3$, and Hermitian matrices, which later were given his name,

$$\sigma_1 = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} & -i \\ i & \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$$

Pauli proposed to represent the electron spin by the three component operator

$$\sigma = \frac{1}{2}\hbar(\sigma_1, \sigma_2, \sigma_3), \quad \hbar = \frac{h}{2\pi}$$

Like Heisenberg, Pauli did not think in terms of group representations at that time; he constructed his two valued wave functions from the Klein-Sommerfeld theory of the spinning top and the complex representation of rotations by Cayley-angles. That was an ingenious and mathematically momentous move towards what little later turned into (Euclidean or relativistic) spinors, although Pauli's hopes to come to a direct explanation of the fine structure of the hydrogen spectrum were not fullfilled at the time.[8] Even the first attempts in 1926 and 1927 to take relativistic effects into account, spinless (Klein-Gordon) or with spin (Darwin), were no more successful in this respect.[9] The situation changed completely in January and February 1928 when Dirac proposed to use 4-component complex-valued "wave" functions $\psi(x) = (\psi_1(x), \psi_2(x), \psi_3(x), \psi_4(x))$ (x in Minkowski-space $\mathbb{M}$) in two successive publications [10]. The ψ- function had to obey the (*Dirac*) equation

$$i\hbar \sum_{\alpha=0}^{3} \gamma^\alpha \frac{\partial}{\partial x^\alpha}\psi = m_0 c\psi \tag{1.13}$$

with (Dirac) matrices γ^μ satisfying the relations $\gamma^j\gamma^k + \gamma^k\gamma^j = \delta_{jk}$ and expressible, e.g., in the form

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & \\ & -\mathbb{1} \end{pmatrix}, \quad \gamma^j = \begin{pmatrix} & \sigma_j \\ -\sigma_j & \end{pmatrix}, \quad 1 \le j \le 3,$$

[6] For a general picture see [31], [30] and [22].
[7] [26]
[8] [30].
[9] [23], [26].
[10] [6].

with (2×2)-unity matrix $\mathbb{1}$ and Pauli matrices σ_j.[11] Thus things looked quite different for Weyl in the late 1920s from what they had been at the end of his first phase of activity in mathematical physics early in the decade. Already in late 1920 he had lost confidence in theories of matter by unification of classical fields according to the Hilbert/Lie approach, including his own one built upon length gauge.[12] While expecting new insights from the rising quantum mechanics, he concentrated on more conceptual or purely mathematical research fields : the analysis of the space problem about 1922/23 and representation theory of Lie groups during the years 1924 to 1926.[13] Weyl kept well informed on the ongoing development during the crucial years for quantum mechanics in the middle of the decade, drawing upon his close scientific relationship with Pauli (1924 – 1928 at Hamburg university), dating from their cooperation on unified geometrical field theories at the turn to and in the early 1920s. Moreover he had contacts to E. Schrödinger who taught at the university Zürich between 1921 and 1927. He apparently felt challenged to contribute to the conceptual and mathematical clarification of the framework of the "new" quantum mechanics, in particular from the point of view of *unitary geometry* (Weyl's title for the first part of his lecture in 1927/28) and the use of *representation theory* of (Euclidean) rotations and permutation for atomic line spectra, Pauli's non-relativistic spin, and mechanism of molecular binding forces.

In winter 1927/28 Weyl had a chance to take up the challenge. Both theoretical physicists working at Zürich had accepted outside calls and had left (P. Debye changed from the ETH to the university Leipzig and E. Schrödinger from the local university to Berlin). Weyl decided to change the subject of a lecture course initially planned and announced on (pure) group theory to one on *Gruppentheorie und Quantenmechanik : Theory of Groups and Quantum Mechanics*. Notes were taken by his assistant F. Bohnenblust and published, after revision and extension, in August 1928 as a book [44], which in the sequel will be abbreviated as *GQM*. In this second book on mathematical physics, Weyl was more cautious than at the occasion of his first one, *Raum - Zeit - Materie* [41] in his expectations of how his contributions might be received by the workers in the field. In the preface to the new book, he remarked :

> It is the second time that I dare to turn up with a book which belongs only partly to my own speciality, mathematics, and partly to physics. I just cannot avoid to play the role of a messenger (often undesired, as I have experienced sufficiently clearly) in this drama of mathematics and physics - fertilizing each other in the dark, although from face to face preferring not to recognize and even renouncing each other. [44][14]

Weyl was not alone in this "role of a messenger" as he realized during the preparation of the lecture notes for publication. Other authors started also in 1927 and 1928 to use group representations in quantum mechanics, among them, most importantly from the mathematical point of view, J. von Neumann and E. Wigner. Also on the physical side, things changed rapidly. Dirac published his papers on the relativistic theory of the electron at the end of the winter semester, in January and February 1928. The impact was enormous and were sufficient reason for Weyl to add to his book a whole new passage on Dirac's equation [44].

[11] Dirac used a slightly different presentation of the matrices than the one given in the text. For a detailed investigation of Dirac's work see [23] or [24].

[12] See [36] or [32].

[13] [21].

[14] Not translated in the English edition by H.P.Robertson.

Another remark in his lectures of 1927/28 leads directly to our geometrical topic.[15] Weyl's gauge idea from 1918, originally linked to a length calibration and " infinitesimal length transport" characterized by a 1-form $\varphi = \sum \varphi_i dx^i$ was rephrased in a quantum mechanical setting by E. Schrödinger, still in a length calibration interpretation [34], and after the rise of the "new" quantum mechanics by V. Fock and F. London in the context of Kaluza-Klein theory of quantum mechanics [8, 25]. The core of their respective arguments dealt with "gauging" a wave function $\psi(x)$ by a point-dependent phase factor $e^{i\lambda(x)}$ (with $\lambda \in \mathbb{R}$) to $\tilde{\psi}(x) = e^{i\lambda(x)}\psi(x)$. The differential of the purely imaginary phase factor, used in Weyl's 1918 theory to "gauge-transform" length connections, could now be used to transform electromagnetic potentials φ_j slightly more convincingly Weyl endorsed this recontextualization of his original gauge idea when he discussed the Schrödinger equation in 1927/28. Probably he had read only the papers by Schrödinger and London, which he cited, not Fock's; but London was aware of and built upon [8].[16] He remarked that the Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi \tag{1.14}$$

containing the Hamilton operator

$$H = \frac{1}{2m}\sum p_j^2 + V(x) \tag{1.15}$$

with potential V and momentum operator $p_j = \frac{\hbar}{i}\frac{\partial}{\partial x^j}$ for a chargeless particle is adequately modified by using the covariant derivative ∂_φ with respect to a potential connection $\varphi = (\varphi_j)$, if a charged particle in field of potential φ is considered. Then the momentum operator becomes

$$p_j = \frac{\hbar}{i}(\frac{\partial}{\partial x^j} + \frac{ie}{\hbar}\varphi_j), \quad i = \sqrt{-1}\,, \tag{1.16}$$

and the Hamiltonian of the Schrödinger theory for the motion of a particle of charge e in an electromagnetic field of potential φ results. Weyl observed that now :

> The field laws satisfied by the potentials ψ and φ of the material and the electromagnetic waves are invariant under simultaneous substitution of[44]
>
> $$\psi \text{ by } e^{i\lambda}\psi, \;\; \varphi_\alpha \text{ by } \varphi_\alpha - \frac{\hbar}{e}\frac{\partial\lambda}{\partial x_\alpha}\ldots$$

He commented that this "principle of gauge invariance" was quite analogous to the one he had postulated in 1918 "by speculative reasons to gain a unified theory of gravitation and electromagnetism" and continued :

> But now I believe that the gauge invariance does not couple electricity and gravitation, but rather *electricity and matter* in the mode presented here. How gravitation according to the general theory of relativity can be included is still uncertain. [44]

[15] This passage was published only in the first edition of [44], no longer in the second edition of 1931 and the English translation.

[16] [39].

Thus Weyl proposed more than a technical adaptation of his old gauge idea to the new framework of QP. In classical UFT the goal was to unify force fields as such in a coherently geometrized, often highly speculative, "a priori" manner, and to derive matter structures from them ; here Weyl indicated a new paradigm centering around the search for conceptual and mathematical structures which link forces to matter fields, without reduction of one to the other and with strong input from experimental evidence.
Classical UFT was, of course, still quite alive at that time. In 1928 A. Einstein turned towards "distant parallelism" for his latest approach to unification. He assumed or postulated, that, in addition to the Levi-Civita connection of the Lorentz metric, an integrable, curvature free, orthogonal connection $\Delta^i{}_{jk}$ with torsion ($\Delta^i{}_{jk} = -\Delta^i{}_{kj}$) is given, which he usually described by a globally parallel system of orthogonal frames. With respect to such an additional structure it was meaningful to consider constant, i.e. point independent, rotations. Although Einstein did not intend so, his additional structure allowed to formulate the Dirac equation in the framework of GRT with distant parallelism and stimulated other physicists to do so.

V. Fock and his Leningrad colleague D. Ivanenko started to explore such an approach in a joint paper submitted to *Zeitschrift für Physik* in March 1929.[17] They hoped to find some "bridge" between gravitation and quantum theory.[18] They started with a formal construct of a linear expression in the Dirac matrices, $ds = \sum_j \gamma_j dx^j$, which they tried to interpret as a matrix valued metric form of some new "linear quantum geometry". From that point of view they hoped to find a kinship between Einstein's field of distant parallelism and the new "linear geometry" [13]. During the following months Ivanenko and Fock realized that the linear structure of the new geometry could better be understood as a covariant derivative of the 4-component complex wave functions which they called "semi-vectors", the later spinors.[19] Still they called the geometry they were heading for "géométrie quantique linéaire" [12, 9].[20] V. Fock continued to explore the terrain and realized soon, that the new covariant derivation of spinors had a much closer kinship with a Weylian phase gauge than with Einstein's distant parallelism. He presented his findings in two articles (no longer co-authored by Ivanenko) to *Physikalische Zeitschrift* and *Comptes Rendus* [9, 10].[21] He thus arrived at a theory combining gravitation, Dirac field, and electromagnetism, which overlapped in large parts with what Weyl achieved in early 1929 when he continued research along the lines indicated in *GQM*.

3 Weyl's and Fock's local spinor structure

Weyl had left Zürich in September 1928 for Bologna (ICM) and Princeton where he spent a year as research professor in mathematical physics.[22] There he could continue, among others, his research on the Dirac equation in general relativity. The approach of distant parallelism did not at all appear convincing to him. He considered it to be a completely

[17] March 25, 1929.
[18] For the group of young relativists in Leningrad see [19], for the early involvement in QP [15]. More on Fock in [18].
[19] The terminology of "semi-vectors" was proposed by L. Landau.
[20] [12] was submitted May 22, 1929.
[21] [9] dated June 24, [10] July 5, 1929.
[22] [14].

"artificial" device and looked for a combined structure of GR and Dirac equation from the point of view of "purely infinitesimal" geometry, which now had to be refined and extended in the light of new physical knowledge. In February 1929 Weyl submitted a first sketch of methods and results under the title *Gravitation and the electron* to the *Proceedings of the National Academy of Sciences* [45]. Three months later he delivered a more extended exposition to *Physikalische Zeitschrift* [**?**].[23] At that time he could not know of Fock's parallel work, neither did he know it when he wrote his third paper on the topic in early summer [47].

Fock, on the other hand, got to know of Weyl's new researches [45] only after he finished his own article for *Physikalische Zeitschrift*. He accepted the common mathematical core of their respective approaches, but emphasized the differences from the physical point of view in a postscript [10]. Weyl apparently got to know Fock's work in summer 1929 and was so fond of the common features of their work that he considered it as establishing essentially one and the same theory. He thus referred to it in the preface to the second edition of *GQM* as the "general relativistic formulation of the quantum laws, which have been developped by Mr. V. Fock and the author [Weyl himself]" [44].[24]

Fock and Weyl applied the method of (pseudo-) orthogonal moving frames in Lorentzian space-time M, i.e. they supposed an

$$\text{orthonormal frame of tangent vectors (ONF)}: \quad e(\alpha, x), \;\; 0 \leq \alpha \leq 3,$$

in each point $P \in M$ with coordinates $x = (x^0, \ldots, x^3)$ (depending differentiably on the point). Tangent vectors v at $x \in M$ could thus be represented in components referring to the coordinate basis (ξ^j), or in components with respect to the ONF ($\xi(\alpha)$ in Weyl's notation) :

$$v = \sum_{j=0}^{3} \xi^j \frac{\partial}{\partial x^j} = \sum_{\alpha=0}^{3} \xi(\alpha) e(\alpha, x) \tag{1.17}$$

Besides (differentiable) change of coordinates, changes of the ONF from $e(\alpha, x)$ to $e'(\beta, x)$ $(0 \leq \alpha, \beta \leq 3)$ had also to be taken into account. The latter were given by point-dependent Lorentz-rotations $\vartheta(x)$ represented by matrices (as the ONF's were given in components with respect to a local coordinate system) :

$$\vartheta(x) = (\vartheta^\alpha_\beta) \in SO(1,3)$$

The parallel transport of a frame by the Levi-Civita connection $\Gamma^i{}_{jk}$ could be expressed in terms of "infinitesimal rotations" o depending linearly on infinitesimal displacements $dx = (dx^j)$ in space-time

$$o^\alpha_\beta = \sum_k \omega^\alpha_{\beta k} dx^k \tag{1.18}$$

[23] Submitted, May 8, 1929.

[24] Weyl saw no chance to give an exposition of this theory in the book *GQM*. In the second edition he rephrased, however, his discussion of the representation of the Lorentz group and of the special relativistic Dirac equation, in particular the decomposition of the 4-dimensional spinors into irreducible 2-dimensional representations.

In more recent terminology : By means of the ONF's Fock and Weyl reduced the group of the affine connection $\Gamma^i{}_{jk}$ to the orthogonal group, and characterized parallel transport in M by the resulting orthogonal connection $\omega^\alpha_{\beta k}$.

In the late 1920s this was standard knowledge. The idea of ONF's had already been introduced by Ricci and Levi-Civita in 1900 ; it had been worked out by differential geometers in the 1920s, most prominent among them E. Cartan (in lectures from 1926/27 published as [4]), J.A. Schouten, R. Weitzenböck, L.P. Eisenhart (in monographs 1926 and 1927). Moreover, orthonormal frames played a central role in Einstein's theory of "distant parallelism", from which Fock (and Ivanenko) took the idea.[25] Fock (still in his cooperation with Ivanenko) and Weyl realized that reduction of the Levi-Civita connection to the orthogonal group by the ONF method allowed to introduce covariant differentiation of spinors.[26] Weyl explained clearly that the orthogonal reduction of the connection was necessary in this context, because "Dirac's quantity" ψ

> corresponds to a representation of the orthogonal group which cannot be extended to the group of all linear transformations. The tensor calculus is consequently an unusuable instrument for considerations involving ψ. [45]

For Weyl , this group theoretic consideration was of great importance. In the early 1920s he had analyzed the role of tensors from the point of view of group represention and found out that all irreducible representations of $GL(n,\mathbb{R})$ with a specified permutation symmetry can be characterized by tensors over $\mathbb{R}^n$.[27] In a language closer to pysicists he explained more in detail :

> Vectors and [tensors] are so constructed that the law which defines the transformation of their components from one Cartesian set of axes [ONF] to another can be extended to the most general linear transformation, to an affine set of axes. That is not the case for [the] quantity ψ, however ; this kind of quantity belongs to a representation of the rotation group which cannot be extended to the affine group. [45]

He admitted that the ONF method used by him resembled Einstein's latest appproach in formal aspects, but insisted that this was only a superficial coincidence.

> But here there is no talk of "distant parallelism" ; there is no indication that Nature has availed herself of such an artificial geometry. I am convinced that if there is a physical content in Einstein's latest formal development it must come to light in the present connection.

And he added a reason that went beyond purely mathematical considerations :

> It seems to me that it is now hopeless to seek a unification of gravitation and electricity without taking material waves into account. [45, 219]

Dirac had shown that the *equation* of the free electron expressed in ψ is invariant under Lorentz transformations without asking for the underlying reprentation of the Lorentz group,[28]

[25] In his main article Fock referred, however, also to [7] [10].

[26] For simplicity, I will no longer always add in the sequel Ivanenko to Fock, even in case that concepts appeared already in their joint work.

[27] See [21].

[28] [6], discussed in [23].

but other authors did so immediately later. F. Möglich calculated the complex 4×4-matrices for the "Dirac-quantity" corresponding to a given Lorentz transformation [27], and J. von Neumann discussed the resulting relation

$$\begin{array}{rccc} \Lambda: & SO^+(1,3) & \longrightarrow & GL(4,\mathbb{C}) \\ & o & \longmapsto & \Lambda(o) \end{array}$$

as a "(multivalued!) 4-dimensional representation of the Lorentz group" [40]. Von Neumann emphasized, very much like Weyl , that something essentially new was introduced into mathematical physics :

> The case of a quantity of 4 components which is no 4-vector has never occurred in relativity theory, the Dirac ψ-vector is the first example of this kind. (ibid., translation E.S.)

Thus, immediately after Dirac's publications on the "spinning" electron, theoretically minded authors realized that the new "Dirac quantity" (Weyl), the "ψ-vector" (von Neumann), or the "semi-vector" (Fock, Landau e.a.) was more than just another technical device, but led to a *conceptual innovation* for mathematical physics. Change of reference systems in special relativity ("Cartesian systems of axes" as Weyl would say) by a Lorentz transformation had to be represented by $\Lambda(o)$ in the ψ-space in a way that could not be extended to general linear transformations and thus could not, in a straight-forward manner, be transferred to general relativity.

At the time when Fock and Weyl approached the problem of a general relativistic formulation of the Dirac equation, the young algebraist B.L. van der Waerden established an algebraic calculus for all possible quantities appearing in any representation of the Lorentz group. His contribution was meant as a sort of service to the physicists, stimulated by a question of P. Ehrenfest who had posed the question to design such an algebraic calculus. Van der Waerden picked up the terminology "spinor" from Ehrenfest and gave him a broad audience [38]. In this work he built upon Weyl's exposition of the representation theory of the Lorentz group in *GQM*.

In distinction to other work about 1929, Fock and Weyl admitted point-dependent (Lorentz-)rotations of ONF in space-time, $o(x) \in SO^+(1,3)$, differentiably depending on x, inducing point-dependent transformations $\Lambda(o(x))$ of the spinor space. While Fock immediately headed for the covariant derivation of a spinor ("semi-vector"), Weyl made the underlying invariance idea explicit. He stated for the "laws" which would be characterized by an action principle and by differential equations derived from it :

> The laws shall remain invariant when the axes in the various points P are subjected to arbitrary and independent rotations. [45]

Variational equations were thus required to be invariant under simultaneous transformations

— of vectors/tensors by Lorentz rotations $o(x)$
— and of the spinors under $\Lambda(o(x))$.

In this way, Weyl and Fock introduced and started to study a *local spinor structure* on the underlying space-time manifold M. Both authors *used* local change of coordinates in the spinor space $\Lambda(o(x))$ (the change of trivialization in later language) accompanying a change of ONF's $o(x)$, and Weyl *discussed* its conceptual role quite clearly, although of course not yet applying the terminology of local bundles trivialization.

Weyl did not mention, however, that for a globalization of the procedure the topology of the M might play a role. Such questions of global existence of an ONF (presupposing parallelizability of M), were posed and answered only in the 1930s by the young generation of topologists (E. Stiefel, H. Whitney), apparently stimulated by Einstein's use of (local) "distant parallelism", not by local spinor structures of Fock and Weyl . Global questions for spinor structures were taken up still another generation later and became a research topic only in the 1960s.[29] Weyl , in his 1929 articles, did not even indicate that there might be an open and challenging question in the relationship between spinor structures on M and its topology.

Of immediate interest, for our authors, was the introduction of an "infinitesimal displacement of semi-vectors" (Fock) or the "invariant change $\delta\psi$ on going from the point P to a neighbouring point P' " [45], i.e. in modern terminology the introduction of a connection and parallel transport in a local spinor structure, lifted from the Levi-Civita connection in the underlying Lorentz manifold. In this point the two authors applied slightly different approaches; Weyl's approach was, as one may expect, more conceptual and Fock's more calculational.

Considering two (infinitesimally) "neighbouring" points $P,\ P'$ with coordinates $x = (x^0, \ldots, x^3)$ and $x' = (x'^0, \ldots, x'^3)$ differing by an "infinitesimal displacement" $dx = (dx^0, \ldots, dx^3)$ Weyl argued that parallel displacement of a frame $\{e(\alpha, P)\}$ from P to P' leads to an infinitesimally rotated frame $\{e'(\alpha, P')\}$ described by an infinitesimal rotation $o = \omega(dx)$ with respect to the ONF-system $\{e(\alpha, P')\}$ in P', in slightyl metaphorical notation

$$\{e'(\alpha, P')\} - \{e(\alpha, P)\} = \omega \cdot \{e(\alpha, P')\} \tag{1.19}$$

(compare equation (6)). The representation Λ induces an infinitesimal tranformation dE (Weyl's notation) in $gl(n, \mathbb{C})$, which depends linearly on dx

$$dE = \Lambda(o) = \Lambda\omega(dx)$$

The "differential $\psi(P') - \psi(P)$", i.e. $d\psi = \sum_j \frac{\partial\psi}{\partial x^j} dx^j$, had to be modified accordingly to give the *covariant differential* $\delta\psi$ of ψ [45] [**?**] :

$$\delta\psi = d\psi + dE \cdot \psi \tag{1.20}$$

This conceptually clear description of the covariant differential, had the advantage that in Weyl's discussion Λ could stand for *any representation of the Lorentz group*, not just Dirac's original 4-dimensional one.

Weyl realized of course, as did von Neumann in 1928, that Dirac's representation can be decomposed into two irreducible representations ρ and ρ^+ (which generate all finite dimensional representations of $SL(2, \mathbb{C})$ by tensor products and direct sums). He gave a beautiful geometrical description of the 2-valued inverse of the covering map[30]

$$SL(2, \mathbb{C}) \longrightarrow SO^+(1, 3)$$

[29] See P. Bourguignon's contribution, this volume.
[30] [**?**].

and took ρ as the identical representation of $SL(2,\mathbb{C})$ and $\rho^+ = {}^t\bar{\rho}$ its adjoint. Then he could write Dirac's representation (up to a permutation of ψ-coordinates) as

$$\Lambda \cong \rho \oplus \rho^+, \tag{1.21}$$

and wrote the 4-spinors (after a linear transformation) as $(\psi_1^+, \psi_2^+, \psi_1^-, \psi_2^-)$.
Fock analyzed the condition (incorporated by Dirac into his new symbolic game) that the ψ-functions get their physical meaning from the condition that the evaluation map

$$\psi \longmapsto (a^0, \ldots, a^3) \text{ with } a^j =< \gamma^j\psi, \psi >, \ 0 \le j \le 3,$$

leads to a vector (a^j). Therefore it was natural to postulate that "changes of a semi-vector ψ under an infinitesimal parallel displacement" are compatible with parallel displacement of vectors. This allowed him to compute matrices $C_l \in GL(4,\mathbb{C})$ which describe such compatible "infinitesimal changes of semi-vectors" (the parallel displacement in the local spinor structure). In his own representation $\tilde{\gamma}^j$ of the Dirac matrices Fock derived the condition

$$C_l = \frac{1}{4}\sum_{j,k,l} \tilde{\gamma}_j\tilde{\gamma}^k\omega^j{}_{kl} + i\phi_l \ , \ \text{ with } \ \tilde{\gamma}_j = \sum_k \epsilon_{jk}\tilde{\gamma}^k, \tag{1.22}$$

$\epsilon = \mathrm{diag}(1,-1,-1,-1)$ the signature diagonal matrix, ω the orthogonally reduced Levi-Civita connection, and ϕ_l any matrix "proportional" to unity

$$\phi_l = f_l\mathbb{1} \quad \text{ with } f_l \text{ real-valued function} \tag{1.23}$$

[10]. Fock thus arrrived at an explicit form of Weyl's infinitesimal spinor transformation dE, at least for the case of the (original) Dirac representation,

$$dE \cdot \psi = \sum_l C_l dx^l\psi$$

On that basis Fock easily expressed covariant differentiation of a spinor with respect to a vector direction of a the frame $\{e(\alpha)\}$

$$D'_\alpha\psi = \frac{\partial}{\partial e(\alpha)}\psi - C_\alpha\psi \tag{1.24}$$

or a coordinate direction x^j

$$D_j\psi = \frac{\partial}{\partial x^j}\psi - \tilde{C}_j\psi \tag{1.25}$$

where $\tilde{C}_j$ are slightly different matrices calculated from the C_j's. For Weyl , both versions of covaraint differentiation could be derived from his "covariant differential" $\delta\psi$ of equation (8).

4 An additional $U(1)$-gauge

Until here I omitted an important observation made by both authors, which led back to Weyl's gauge idea. The "lifting" of the Levi-Civita connection to the spinor structure was

not uniquely determined, even if we neglect the double valuedness of the $SL(2,\mathbb{C})$ covering of the Lorentz group.

Fock's calculation of the matrices (equation (10)) showed that the compatibility condition determines the C_l only up to addition of purely imaginary matrices $if_l\mathbb{1}$. Covariant differentiation of spinors (equations (12), (13)) is then affected by an additive term $-if_\alpha\psi$. In a kind of *déja vu* Fock realized that the additional term could be percieved as derived from a phase-gauge factor of the ψ-field :

> The appearance of the Weylian differential form in the law of parallel displacement stands in close relation to the fact remarked by the author [Fock] and also by Weyl (...) that the addition of a gradient to the 4-potential corresponds to a multiplication of the ψ-function by a factor of absolute value 1. [10]

On that basis, Fock formulated the Dirac equation for the general relativistic electron by covariant derivation in his local spinor structure, including a Weylian $U(1)$-gauge term as an integrated part of the covariant derivation (13) (ibid.)

$$F\psi = 0 \quad \text{with} \quad F = i\hbar\sum_{j=0}^{3}\gamma^j D_j + mc\gamma_4 \tag{1.26}$$

Weyl discussed the question similarly. although slightly more general. He argued that any semantically relevant information derived from a spinor field was invariant under $U(1)$-symmetries of the spinor representation, because the $SO^+(1,3)$-covariants used to represent physical quantities were given by Hermitian forms $<\psi, A\psi>$ and thus invariant under multiplication by a phase factor $e^{i\lambda}$ of ψ. Therefore the spinor connection ("the infinitesimal linear transformation dE of the ψ") is determined by the "infinitesimal rotations" ω of the reduced Levi-Civita connection only up to "a purely imaginary multiple $i \cdot df$ of the unit matrix". In other words, with dE also

$$dE' = dE + idf\mathbb{1}$$

is compatible with the underlying metric of GRT. Weyl concluded :

> For the unique determination of the covariant differential $\delta\psi$ of ψ such a df for each line element $\overrightarrow{PP'} = (dx)$ starting from P is needed. [**?**]

The selection among the spinor connections compatible with the Levi-Civita connection could justly be considered as a "gauge", in strong analogy to the length gauge of 1918. Morover Weyl used, just like Fock, the possibility to express the Dirac equation of the electron in an electromagnetic field by means of covariant differentiation of spinors *including* a $U(1)$-gauge potential ("such a df").
For action functions applying to spinor fields he felt it legitimate to postulate :

> If one (...) substitutes
>
> $$\psi \text{ by } \quad e^{i\lambda} \qquad f_p \text{ by } \quad f_p - \frac{\partial\lambda}{\partial x_p}$$
>
> with λ an arbitrary function of the position, gauge invariance necessarily holds, in the sense that the action principle remains invariant. [**?**]

From the point of view of infinitesimal symmetries, the new gauge structure resembled in certain features Weyl's study of the *Raumproblem* early in the 1920s. In the analysis of the space problem he had characterized "congruences" by a subgroup G of $SL(n,\mathbb{R})$, contained in a larger group H of "similarities", in which G was normal (in fact, H was the normalizer of G in $GL(n,\mathbb{R})$). One of his postulates was a uniqueness condition for affine connection equivalent (in a certain sense) to a given linear connection in the larger group. In 1929 he again dealt with a pair of groups, now given by physical considerations, the smaller one being the Lorentz group or its universal convering, $G = SL(2,\mathbb{C})$, and the larger one was $H = SL(2,\mathbb{C}) \times U(1)$ in which G was normal by construction. Again a uniqueness condition for a connection, compatible to another given one, played a crucial role for the analysis. The uniqueness condition was now formulated "bottom up", i.e. from a given (Levi-Civita) connection in the smaller group to the larger one, and uniqueness of the (spinor) connection with respect to the larger group was achieved only by adding a connection in the qotient group $U(1)$ (respectively bundle, from the later point of view). In this sense there was a structural analogy considering group extensions for infinitesimal symmetries, although the methodology had changed considerably. In 1929 Weyl no longer tried to found his approach on a priori principles, but rather analyzed symbolic forms worked out ("constructed") by mathematical physicists in close communication with experimental knowledge of the rising quantum physics.

Weyl discussed how one could arrive at physical consequences from his approach. It would lead us to far to follow this line here.[31] I just want to mention that Weyl drew impressive consequences from the postulate of invariance of the action integral under infinitesimal symmetries of different kinds :

— infinitesimal rotations of the frames leads to symmetry of the energy-momentum tensor,

— infinitesimal coordinate translations leads to "quasi"-conservation of energy and momentum and in the case of special relativity by integration to invariance of rotational momentum [**?**],[32]

— infinitesimal $U(1)$ gauge transformations leads to conservation of charge (ibid., 264f.).

He hoped, morover, that his general relativistic approach to the Dirac equation, together with the separation of the spinor fields into components of irreducible representations ρ and ρ^+ might lead to a solution of the problem of *negative energies* in the original Dirac equation. In late 1929 Dirac proposed a solution to this problem by some imaginative ad-hoc arguments postulating the existence of positive electrons (positrons) appearing as constitutive parts of the solution of the original Dirac equation with non-vanishing mass term, and surprising "fluctuations" between positive and negative charge contributions to it. It turned out that neither the positive charge contributions could be separated nor the resulting "fluctuations" eliminated from the solution [24].

[31] Cf. [37].

[32] Weyl spoke of "quasi-conservation" of energy-momentum t_p^q, because of a second term in the differential equation derived from invariance under infinitesimal translations :

$$\frac{\partial t_p^q}{\partial x_q} + \frac{\partial e^q(\alpha)}{\partial x_q} t_q(\alpha) = 0$$

Literal conservation of energy and momentum holds only, if the respective terms of the gravitational fields are added or, in special relativity, after specialization of the ONF's [**?**].

Weyl , for his part, attempted for a short while in 1929 to avoid such fluctuations by the proposal to study solutions of a modified Dirac equation in the irreducible components of the representation ρ and ρ^+ separately (Weyl spinors). He remarked, however, that in this equation no mass term could be included without losing gauge invariance [45]. As a research strategy to overcome the problem he proposed to neglect at first, on the level of the spinor equation, the mass of the electron and to reconstruct it, in a second step of theory development, as an integral invariant that couples to gravitation.

> Be bold enough to leave the term involving mass entirely out of the field equations. But the integral of the total energy density over space yields an invariant, and at the same time constant, mass ; *require of it that its value be an absolute constant of nature m* which cannot vary in value from case to case. This introduction of mass is born of the idea that the inertia of matter is due to its energy content. [45]

Such an approach made sense only in a joint theory of gravitation, quantum physics (in the sense of the modified Dirac equation) and electromagnetism. In his attempt for an integrated theory Weyl now pursued the concrete goal to contribute to the solution of the mass problem of the electron.

The proposal to start from a "massless" electron was rejected by physicists immediately. In the postscript to his article for the *Physikalische Zeitschrift*, Fock argued strikingly (and presumably also convincingly for Weyl)[33] that the current of the Weyl-spinor field was lying on the light-cone. Thus there remained no realistic hope for a solution of the electron's mass problem along the line indicated by Weyl [10]. Similarly Pauli rejected Weyl's proposal to circumvent the mass problem for the electron, although from a conceptual point of view he found the new integration of the gauge idea into quantum physics most convincing. He contributed essentially to its dissemination and survival in the physics community. Moreover he revived Weyl spinors in 1956 when he looked for an adequate mathematical representation of his newest hypothetical entity, the neutrino. This is a different and historically complicated story which cannot be dealt with here.[34]

Weyl indicated that field quantization was another problem that had to be solved before one might hope for an answer to the questions raised :

> Another difficulty which stands in the way of a comparison with experience is that the field equations must first be quantized before they can be applied as a basis for the statistics of quantum transitions. But our theory is also hopeful in this rerspect inasmuch as the anti-symmetric Fermi statistics of the electrons, corresponding to the Pauli exclusion principle, here necessarily leads to the symmetric Bose-Einstein statistics of photons. [45]

Weyl could probably not surmise which tremendous difficulties had to be surmounted on the path indicated here. When he reworked *GQM* for the second edition he knew already more about the nature of problems arising from the infinities of field quantization. He made some striking observations with respect to symmetries in quantum electrodynamics, but did

[33] In the 2nd edition for *GQM* Weyl no longer insisted on his 1929 proposal and supported Dirac's strategy to deal with the problem [44].

[34] See [30], [37].

not contribute to its further development in the later 1930s and 40s.[35]

5 Geometry and physics : interpretations and perspectives

As we have seen, Weyl's and Fock's 1929 work contained a strong common mathematical core. They both established local spinor structures on Lorentz manifolds with an additional internal $U(1)$ symmetry and proposed to use a connection in this structure, determined by or determining gravitation and electromagnetism and governing the motion of the spinor field. But they had strong differences with respect to the question of how geometry and physics could or should be related.

Fock proclaimed that his goal was "the geometrization of Dirac's electron theory and its subsumtion (Einordnung) in general relativity" [10]. This was a conceptual-methodological task, rather than one of concrete physical theory building. He hoped, however, that his investigation might "contribute to the solution of the problems" in Dirac's theory, referring apparently to the paradox of negative energies and positive probability of fluctuations between negative and positive energies, respectively charges. He thus expected that his geometrization of the Dirac operator might lead, in the long run, to a progress of physical theory in a more technical sense. Fock's main hope was, however, to contribute to what he (and Ivanenko) thought to be a challenging goal of contemporary physics, the development of a common conceptual structure for relativity and quantum physics.

V. Fock had learned relativity from A. Friedmann and participated prominently in the development of relativity theory in Russia.[36] In the later 1920s he stood in close contact to a group of young physicists in Leningrad around L. Landau, G. Gamow, and M. Bronstein, to which his early 1929 coauthor D. Ivanenko belonged. The young physicists enthusiastically supported the cultural awakening in the early Soviet Union and wanted to contribute to it by their work in relativity and quantum physics.[37] This was apparently part of the background for Fock's and Ivanenko's premature claim to have found a path towards *quantum geometry.*

In a letter to *Nature*, dated March 21, 1929, they announced their first, still very sketchy ideas on "linear geometry" as a contribution to this challenging task.[38] In the *Comptes Rendus* note of May 22, 1929, they shifted attention in their "géométrie quantique linéaire" from the "matrix valued linear metric" to parallel displacements and covariant differentiation in a local spinor structure. Once more, they claimed to have found a method to reconcile quantum physics with geometry

> Il importe de signaler un point qui distingue les idées exposées dans cette Note de celles d'Einstein et de Levi-Cività : c'est l'intervention des matrices-opérateurs

[35] For Weyl's contribution to the symmetries in early quantum electrodynamics, see [5] ; for the history of quantum electrodynamics [35].

[36] [19].

[37] [15].

[38] With respect to their purely formal "linear form with matrix coefficients" $ds = \sum_k \gamma_k dx_k$ (see above) they proclaimed : "This linear ds is connected with Dirac's wave equation in the same way as the Riemannian ds^2 with the relativistic wave equation of the older theory. This linear geometry seems to furnish a basis on which a uniform theory of gravitation, radiation, and quantum phenomena is to be constructed" [11]. For more details they referred to their forthcoming paper [13]

> dans les équations pour les quantités purement géométriques. Grâce à cela on peut bien s'imaginer un champ électromagnétique dans un espace euclidien, ce qui était impossible dans les autres théories. [12]

In his later contributions Fock was more cautious and weakened the claim to the more moderate one of having pursued "the geometrization of Dirac's theory of the electron and its subsumtion under the general theory of relativity" [10]. He admitted that the "difficulties which are inherent in Dirac's theory" had not yet been touched, but added :

> Our investigations might perhaps contribute indirectly to the solution of these difficulties, by showing what the original unchanged Dirac theory can achieve. (ibid.)

The reference to the "original unchanged Dirac theory" was probably formulated after Fock got to know Weyl's proposal and indicated a disassociation from the latter, the reasons of which were explained in the postscript. Fock thus proclaimed that the geometrization of the Dirac equation by the spinor structure with connections and covariant derivation was an important methodological achievement in itself.
In this point Weyl did not agree at all. He had lost confidence in the geometrical unification programs which he himself had contributed so effectively by his gauge unification in 1918. About the end of the 1920s he no longer expected any deeper understanding of physical reality by the still blossoming geometrical unification programs.[39] He criticized, in particular, Einstein's latest attempt of unification by an additional structure of distant parallelism as a turn towards a physically unmotivated "artificial geometry" [45]. In his later 1929 paper for *Physikalische Zeitschrift* he argued more in detail :

> I am unable to believe in distant parallelism by several reasons. Firstly, a priori, my mathematical sense (mathematisches Gefühl) opposes against accepting such an artificial geometry ; for me, it is difficult to conceive of a power which would make the local systems of axes, in their twisted position in the different world-points, freeze together in rigid affiliation. Moreover, two important physical reasons have to be added [**?**].

As "first physical reason", Weyl mentioned his gauge theory of electromagnetism. He argued that only the point-dependence of the ONF's gave rise to a variable phase factor $e^{i\lambda}$ and thus the new principle of gauge invariance. The "second physical reason" was, to Weyl , the possibility to derive symmetry of the energy-momentum tensor and the invariance of rotational momentum in special relativity from infinitesimal rotations of the ONF's or of infinitesimal translations of coordinates (see above). Thus Weyl's "physical reasons" consisted essentially of methodological arguments for the superiority of invariance properties in an infinitesimal symmetry approach, close to those which about three decades later became central in the rise to prominence of more general "gauge" theories.[40]

The 1930 Rouse Ball lecture at Cambridge university gave Weyl the opportunity to explain his view of the unification programs to a wider scientific audience. He still considered the attempts "to geometrize the whole of physics", undertaken after Einstein had so successfully geometrized gravitation, very comprehensible at its time [48]. He explained his own

[39] On the "diversity" of these programs see [17].
[40] Cf. [28].

theory of 1918 and resumed its critical reception by physicists. He reviewed Eddington's approach to unification by affine connections and Einstein's later suppport for that subprogram, always in comparison with his own "metrical" unification of 1918, and concluded that in hindsight one could see that both theory types were "merely geometrical outfits (geometrische Einkleidungen) rather than proper geometrical theories of electricity". He ironically added that the struggle between the metrical and affine UFT's (i.e. Weyl 1918 versus Eddington/Einstein) had lost importance, as in 1930 it could no longer be the question which of the theories would "prevail in life", but only "whether the two twin brothers had to be buried in the same grave or in two different graves" (ibid., 343). He again made clear that he could not find any argument in favour of Einstein's distant parallelism approach, nor could he find good prospects for the Kaluza-Klein approach.[41] Weyl even accused Einstein's new theory of "breaking with the infinitesimal point of view. [...] The result is to give away nearly all what has been gained in the transition from special to general relativity. The loss is not compensated by any concrete gain" [48].
Weyl percieved a nearly complete scientific devaluation of the UFT's of the 1920s, resulting from developments in the second part of the decade :

> In my opinion the whole situation has changed during the last 4 or 5 years by the detection of the matter field. All these geometrical leaps (geometrische Luftsprünge) have been premature, we now return to the solid ground of physical facts. [48]

He continued to sketch the theory of spinor fields, their phase gauge and its inclusion into the framework of general relativity along the lines of the 1929 articles. Weyl emphasized that, different from the principles on which the classical UFT's had been built, the new principle of phase gauge "has grown from experience and resumes a huge treasury of experimental facts from spectroscopy" (ibid. 344). He still longed for safety, just as much as at the time after the First World War, when he designed his first gauge unification. Now he no longer expected to achieve it by geometric speculation, but tried to anchor it in more solid grounds :

> By the new gauge invariance the *electromagnetic field now becomes a necessary appendix of the matter field, as it had been attached to gravitation in the old theory.* [48]

Weyl made it very clear to his readers that he had changed his perspective. He no longer saw a chance in attempts to derive matter in highly speculative approaches from mathematical structures devised to geometrize force fields ; he now set out to search forms for the mathematical represention of matter, which gave expression to the enduring traces in the "huge treasury" of experimental knowledge. For him, this was reason enough to prefer the view that the electrical field "follows the ship of matter as a wake, rather than gravitation" (ibid.).

In short, *Weyl had turned from his idealist approach to matter*, pursued at the turn to the 1920s, *to a symbolic realist* one at the end of the decade. This change of perspective had

[41] The revival of Kaluza-Klein type theories in the 1980s happened in a completely different content of theory development. In this conference, moreover, P. Cartier argued that there are reasons which might lead to a renewed interest in the original form of Weyl's *purely infinitesimal geometry* — again in a modified physical interpretation and theory context.

consequences for his views on geometrization. With reference to Fock's interpretation of the role of geometry in the general relativistic Dirac equation Weyl continued :

> Mr. Fock calls the derivation of the new gauge invariance from general relativity, which he arrived at nearly simultaneously with me, a geometrization of Dirac's theory of the electron. In this respect I cannot agree with him. My impression is that we have abandoned geometrization by linking electricity to matter rather than to gravitation. I fear that the geometrizing tendency, which seized gravitation in full right and supported by the most intuitive arguments, was misled when it was extended to other physical entities. [48]

Weyl did not, on the other hand, completely negate any *possibility* to find a geometrical quantum theory. He only warned that, if one wanted to continue with the geometrizing tendency, one had to invent a "natural geometry" leading to a spinor type field ψ for the characterization of its structure, in addition to the ONF. Whereas Fock claimed to have achieved this already, Weyl remained agnostic :

> One had to set out in search of a geometrization of the matter field ; if one succeeds here, the electromagnetic field is added as a premium to the bargain. I have no idea what kind of geometry this might be. (ibid.)

From the perspective of late 20th century developments in differential geometry and the tremendous role of gauge field theories, Weyl's evaluation is highly surprising and even seems paradoxical : Why did he not percieve his own and Fock's invention of local spinor structures with additional $U(1)$-gauge as a sufficiently rich extension of geometry to deal with matter structures ?[42]

Our own perspective has been shaped by the development of differential geometry and topology in the second half of the last century, which was deeply influenced by Elie Cartan's work, the work of his students and other researchers. In the late 1950s and 1960s bundle structures with their inbuilt transformation behaviour have become central concepts in geometry and topology. In this sense, Weyl's first desideratum of a "natural geometry" which includes spinor type field in its core structure seems to be satisfied, and it becomes difficult to grasp why Weyl, unlike Fock, did *not* accept their common contribution as a valuable step in this direction.

We may assume that Weyl over-emphasized his scepticism with respect to geometrization of physics at the turn to the 1930s, because he still wanted to correct his earlier exuberance in this respect. Moreover he wanted to disassociate himself strongly from the "old" unification programs which where still alive in the latest attempts of Einstein, or Kaluza and Klein, and wanted to counteract them in the scientific discourse as clearly as possible.

For a proper historical understanding we have to take another aspect into account. Weyl's attempts to integrate geometry with physics had, from their very beginnings after the First World War, a strong intentional reference to the quantum stochastical aspects of matter as a "dynamical agens", even at a time when these were not understood at all. In the early 1920s Weyl had dared to speculate in wide leaps about a possible relationship between the intuitive, the mathematical and the physical understanding of the continuum, some inbuilt discrete "free-choice" structures and the end of classical determinism in natural science.[43] In

[42] I thank Jim Ritter who indicated this point to me and insisted on a closer historical perspective.

[43] Most prominent and controversially discussed in this respect is [**?**].

1925, in his manuscript for the Lobachevsky centenary volume (published only posthumously [43]), Weyl indicated that the vagueness of physical determination of space-time localization has to be taken serious for the basic theoretical structure of geometry. This vagueness ought to be considered a principal feature for the mathematical characterization of geometry and to be dealt with, in principle, in some stochastical approach informed by "the actual state of physics", i.e. quantum physics. But then, so Weyl remarked, at a time when the "new" quantum mechanics was just being shaped, the question, how such a quantum stochastical foundation for geometry relates to the differentiable structure of classical geometry, turned into a completely open problem. He ended the passage by the honest remark :

> One has to admit that until now nearly nothing has been achieved for the question what it means to apply differential calculus to [physical] reality. [43]

With such questions Weyl was not *completely* alone. But they were far from what most physicists or mathematicians considered useful at the time, or even later in the 1930, when Fock's young colleague M. Bronstein explored the questions of necessary revision of time-space concepts from the point of view of quantum physics [15]. Fock's hope of 1929 to leave classical geometry behind and to turn towards geometrical quantum structures was comparably innocent. With such a point of view he was content with an extension of differential geometry which would appear, at most, as a semi-classical enrichment.
In his 1930 talk at Cambridge (and the later publication) Weyl expressed clearly that from a proper *geometry of matter* he expected a deep break with the classical tendency of geometrization prevailing in the UFT's. He was less clear, to say the least, what should be substituted for it ; but there were strong reasons for such vagueness. His own approach to the mass problem of the electron had turned out to be unsatisfactory ; Dirac's alternative appeared more promising, but had still a long way to go before, at least, a technically valid solution of the quantization problem was in sight[44] — not to speak about the extensions of later quantum gauge field theories and the still unanswerable question of the mass spectrum of basic constituents of matter. Therefore Weyl's remark "I have no idea what kind of geometry this might be", was just as honest as his comment in 1925 that "nearly nothing had been achieved" for a semantically reliable relation of the differentiable structure of geometry to the "actual state of physics".

Notwithstanding a whole range of new open questions and desiderata, inlcuding the one for a historical evaluation of recent developments, we now see several candidate programs for a quantum geometry aiming at (or preparing) a unification of quantum field theories. It is not yet clear, whether one of them (or perhaps several) will "prevail in life". Weyl's proposal to look for a "geometry of matter" informed by the treasury of experimental knowledge could still be taken as an advice for a critical discourse in and among the different research programs.[45] Perhaps future developments will show, whether Weyl's guess that the geometrization of interaction and metrical fields is "added as a bonus", once a proper geometry of matter has been achieved, is just another speculative dream. It still may turn out that it indicates a hint for an appropriate theory development.

[44] See [35].

[45] At the turn of the century we may add that, in addition to recent and coming results in high-energy spectroscopy, geometrical aspects of low energy EPR-type experiments constitute a valuable novel part of the "treasury" of experimental knowledge, which ought to be taken into account in a future "geometry of matter".

Bibliographie

[1] L. Boi, D. Flament, J.-M. Salanskis eds. *1830 - 1930 : A Century of Geometry, Epistemology, History and Mathematics*, Berlin etc., Springer, 1992.

[2] T. Yu. Cao, *Conceptual Developments of 20th Century Field Theories*, Cambridge, University Press, 1997.

[3] Cao, Tian Yu ed. *Conceptual Foundations of Quantum Field Theories*, Cambridge, University Press, 1999.

[4] E. Cartan, *Leçons sur la géométrie des espaces de Riemann*, Paris, Gauthier-Villars, 1928.

[5] R. Coleman, H. Korté, *Hermann Weyl : Mathematician, physicist, philosopher*, In *[33]*, pp. 161–388, 2001.

[6] P. A. M. Dirac,
The quantum theory of the electron I, *Proceedings Royal Society London* A 117, 610–624, Collected Works Cambridge 1995 , 303–319, The quantum theory of the electron II, *Proceedings Royal Society London* A 118, 351–361, Collected Works, 321–333, 1928.

[7] L. Eisenhart, *Riemannian Geometry*, Princeton University Press, 1926.

[8] V. Fock, *Über die invariante Form der Wellen- und Bewegungsgleichungen für einen geladenen Massepunkt* , Zeitschrift für Physik 39, 226–232, 1926.

[9] V. Fock, *Sur les équations de Dirac dans la théorie de la relativité générale* , Comptes Rendus Académie des Sciences Paris 188, 25–27, Séance du 24 juin 1929, 1929a.

[10] V. Fock, *Geometrisierung der Diracschen Theorie des Elektrons* , Zeitschrift für Physik 57, 261–277, 1929b.

[11] V. Fock, D. Iwanenko, *Quantum geometry* , Nature 123, 838, Letter March 21, 1929, 1929a.

[12] V. Fock, D. Iwanenko, *Géométrie quantique linéaire et déplacement parallèle*, Comptes Rendus Académie des Sciences Paris 188, 1470–1472, Séance du 22 mai 1929, 1929b.

[13] V. Fock, D. Iwanenko, *Über eine mögliche Deutung der relativistischen Quantentheorie*, Zeitschrift für Physik 54 , 798–802, 1929c.

[14] G. Frei, U. Stammbach, *Hermann Weyl und die Mathematik an der ETH Zürich, 1923 – 1930*, Basel etc., Birkhäuser, 1992.

[15] V. Frenkel, G. Gorelik, *Matvei Petrovich Bronstein and Soviet Theoretical Physics in the Thirties*, Basel etc,, Birkhäuser, 1994.

[16] P. Galison, *Theory bound and unbound : Superstrings and experiments*, In *Laws of Nature, Essays on the Philosophical, Scientific and Historical Dimensions*, ed. Friedel Weinert, Berlin etc., de Gruyter pp. 369–408, 1995.

[17] C. Goldstein, J. Ritter, *The varieties of unity : Sounding unified theories 1920–1930*, Berlin, MPI for History of Science, Preprint 149, 2000.

[18] G. Gorelik, *Vladimir Fock : Philosophy of gravity and gravity of philosophy*, in *The Attraction of Gravitation : New Studies in the History of General Relativity*, ed. John D. Earman John ; Janssen, Michel ; Norton. Basel etc., Birkhäuser, pp. 308–331, 1993.

[19] G. Gorelik, V. Vizgin, *The reception of the theory of relativity in Russia and the USSR*, in *The Comparative Reception of Relativity*, ed. Thomas Glick. Dordrecht, Reidel, pp. 265–326, 1987.

[20] J. Gray ed. , *The Symbolic Universe : Geometry and Physics 1890–1930*. Oxford University Press, 1999.

[21] T. Hawkins *Emergence of the Theory of Lie Groups. An Essay in the History of Mathematics 1869–1926*. Berlin etc., Springer, 2000.

[22] J. Hendry *The Bohr-Pauli Dialogue and the Creation of Quantum Mechanics*. Dordrecht, Reidel, 1984.

[23] H. Kragh, *The genesis of Dirac's relativistic theory of electrons*, Archive for the History of Exact Sciences 24, 31–67, 1981.

[24] H. Kragh *Dirac : A Scientific Biography*. Cambridge University Press, 1990.

[25] F. London *Quantenmechanische Deutung der Theorie von Weyl*, Zeitschrift für Physik 42, 375–389, 1927.

[26] J. Mehra, H. Rechenberg, *The Historical Development of Quantum Theory. Vol 6, The Completion of Quantum Mechanics 1926–1941. Part I*. Berlin etc., Springer, 2000.

[27] F. Möglich, *Zur Quantentheorie des rotierenden Elektrons*, Zeitschrift für Physik 48, 852–867, 1928.

[28] M. Morrison, *The new aspect : Symmetries as meta-laws — structural metaphysics*, in *Laws of Nature. Essays on the Philosophical, Scientific and Historical Dimensions*, ed. Friedel Weinert. Berlin, de Gruyter, pp. 157–188, 1995.

[29] L. O'Raifeartaigh, N. Straumann, *Gauge theory : Historical origins and some modern developments*, Reviews of Modern Physics 72, 1–23, 2000.

[30] A. Pais, *Inward Bound : Of Matter and Forces in the Physical World*. Oxford, Clarendon, 1986.

[31] H. Rechenberg, *Quanta and quantum mechanics*, in *Twentieth Century Physics, vol I*, ed. B. Pippard Laurie Brown, A. Pais. Bristol, Philadelphia, New York,, Institute of Physics Publishing chapter 3, pp. 143–248, 1995.

[32] E. Scholz, *Weyls Infinitesimalgeometrie, 1917–1925*, In [33]. pp. 48–104. 2001*a*.

[33] E. Scholz, *Hermann Weyl's Raum - Zeit - Materie and a General Introduction to His Scientific Work*, Basel etc., Birkhäuser. 2001*b*.

[34] E. Schrödinger, *Über eine bemerkenswerte Eigenschaft der Quantenbahnen eines einzelnen Elektrons*, Zeitschrift fúr Physik, **12**, 13–23, 1922.

[35] S. S. Schweber, *QED and the Men Who Made It : Dyson, Feynman, Schwinger, and Tomonaga*. Princeton, University Press. 1994.

[36] S. Sigurdsson, *Hermann Weyl, Mathematics and Physice, 1900 – 1927*, Cambridge, Mass., PhD Dissertation, Harvard University. 1991.

[37] N. Straumann, *Ursprünge der Eichtheorien*, In [33], pp. 138–160, 2001.

[38] B. L. van der Waerden, *Spinoranalyse*, Nachrichten Gesellschaft der Wissenschaften zu Göttigen, pp. 100–109, 1929.

[39] V. Vizgin, *Unified Field Theories in the First Third of the 20th Century*, Translated from the Russian by J. B. Barbour, Basel etc., Birkhäuser, 1994.

[40] J. von Neumann, *Einige Bemerkungen zur Diracschen Theorie des relativistischen Drehelektrons*, Zeitschrift für Physik, **48**, 868–881, 1929.

[41] H. Weyl, *Raum, - Zeit - Materie*, Berlin etc., Springer 1918. Further editions, 1919, 1919,1921, 1923, 1970, 1988.

[42] H. Weyl, *Das Verhältnis der kausalen zur statistischen Betrachtungsweise in der Physik*, Schweizerische Medizinische Wochenschrift, p. 10p, GA II, 113–122, 1920.

[43] H. Weyl, *Riemanns geometrische Ideen, ihre Auswirkungen und ihre Verknüpfung mit der Gruppentheorie*. Berlin etc., Springer, 1925/1988.

[44] H. Weyl, *Gruppentheorie und Quantenmechanik*, Leipzig, Hirzel, 2ed 1931, English translation by H.P. Robertson 1931. Abbreviated as *GQM* , 1928.

[45] H. Weyl, *Gravitation and the electron*, Proceedings of the National Academy of Sciences, USA, **15** , 323–334. GA III, 217–228, 1929*a*.

[46] H. Weyl, *Elektron und Gravitation*, Zeitschrift für Physik, **56**, 330–352. GA III, 245–267, 1929*b*.

[47] H. Weyl, *Gravitation and the electron*, The Rice Institute Pamphlet, **16**, 280–295. GA III, 229–244, 1929*c*.

[48] H. Weyl, *Geometrie und Physik*, Die Naturwissenschaften, **19**, 49–58. Rouse Ball Lecture Cambridge, May 1930. GA III, 336–345, 1931.

Les groupes comme Universaux

Jean-Marie Souriau

L'histoire des sciences respecte scrupuleusement la chronologie des idées. L'un des rôles de l'épistémologie, c'est de déchirer ce livre d'histoire, et d'en ajuster autrement les feuillets.

Plaçons-nous pour commencer au $14^{\text{è}me}$ siècle, et lisons Guillaume d'Ockham :
"*Un Universel, c'est un projet de l'esprit qui s'applique à un grand nombre d'objets*".

L'Universel groupe remonte aux années 1870 (Sophus Lie, Felix Klein). Il a d'abord été "**groupe de transformations**" (Lagrange 1770, Abel 1826, Galois 1832). Puis il est devenu "**abstrait**", et accompagné par les "**actions de groupe**", qui l'appliquent à un grand nombre d'objets.

Dirons-nous aujourd'hui, comme Felix Klein : "une géométrie, c'est une action de groupe" ? Non, choisissons plutôt :

une géométrie, c'est un groupe

1. Commençons par la

géométrie euclidienne

construite au $III^{\text{è}me}$ siècle en utilisant les déplacements des **"points de l'espace"**. Déplacements qui constituent à l'évidence un "**groupe de permutations**", le

Groupe d'Euclide

Le Groupe d'Euclide "**abstrait**" peut aussi s'appliquer à d'autres objets que les points. Euclide l'applique aux objets *couples-de-points*, puis aux *triplets-de-points*, etc.

Ces divers objets sont classés par l'action du groupe : deux couples de points sont "de même classe" si un déplacement amène l'un sur l'autre. Les classes de couples, ce sont donc

les "distances". Ce qui pose immédiatement le problème du "troisième cas d'égalité des triangles".

Le Groupe d'Euclide fait aussi apparaître des objets *réguliers* (polyèdres par exemple) ; leur *régularité*, c'est le sous-groupe constitué des éléments qui ne changent pas l'objet.

Ces notions duales : *classe* et *régularité*, s'étendent évidemment *à toute action de tout groupe*. Les classes d'une action de groupe peuvent s'interpréter "*en compréhension*" ou "*en extension*" — comme l'enseigne la *Logique de Port-Royal* (1662).

2. Le Groupe d'Euclide, nous venons de le définir par son action *sur les points*. Mais notre programme consiste à attribuer la primauté *au groupe*, indépendamment du choix d'une action particulière :

au revoir et merci, chers petits points.

Au $III^{\text{ème}}$ siècle déjà, Archimède avait utilisé des objets euclidiens qui ne sont pas constitués de points :

les moments

Un moment peut se définir par son "**travail**" dans un "**déplacement virtuel infinitésimal**" (Lagrange , 1788). Autrement dit, un moment est un "tenseur" défini sur le groupe d'Euclide, au point neutre.

La notion de *moment* s'applique ainsi aux groupes qui possèdent une *structure infinitésimale* convenable (les groupes "**difféologiques**", "**groupes de Lie**" par exemple). Alors le groupe agit naturellement sur ses moments : "**action coadjointe**" dit-on.

La statique selon Archimède (*leviers, corps flottants*) s'exprime ainsi : si un corps est en équilibre, la somme des moments qu'il subit est nulle.

Ce "principe d'Archimède" implique donc *les moments et l'action coadjointe* dans la *physique de la matière*.

Un intéressant théorème du $20^{\text{ième}}$ siècle : la géométrie induite sur chaque "**classe de moments**" s'accompagne d'une "**sur-géométrie**", dite "**symplectique**".

Prenons l'exemple d'un moment euclidien réduit à une seule force — comme la poussée de l'eau sur un navire au repos.

La "**classe coadjointe**" de cette force est donc *symplectique* ; plus précisément, c'est une "**variété symplectique de dimension 4**". On peut la représenter comme *l'ensemble des droites orientées*, transmettant chacune une force de la même intensité que la force choisie initialement.

Nous rencontrons dans la nature, des **droites orientées** : les *rayons lumineux*.

La *géométrie euclidienne de la lumière* s'accompagne donc nécessairement d'une *géométrie symplectique*. Essayons de l'analyser.

- Chaque classe euclidienne de rayons lumineux se caractérise par une **"grandeur"** (qui correspond à l'intensité de la force) : cette grandeur, c'est la *couleur*.

- Quand un rayon lumineux "**monochromatique"** traverse un instrument d'optique, il en ressort avec la même couleur, et la correspondance [entrée. sortie] respecte la "**forme symplectique**". *A*ussi bien pour les "**systèmes dioptriques"** (les verres de lunettes) que pour les "**systèmes catadioptriques"** (les miroirs ardents de Syracuse ou les télescopes).
- Un petit problème : il existe un phénomène gênant, appelé **"diffraction"**, qui empêche d'isoler un rayon lumineux. La diffraction permet, au mieux, de concentrer la lumière sur un ensemble de rayons appelé **"faisceau lumineux"** ; et pour le géomètre, ce faisceau est une **"variété lagrangienne"** (dont la définition est **symplectique**). Conséquence : un faisceau lumineux **"lagrangien"** est encore lagrangien quand il ressort d'un instrument d'optique. Voilà pourquoi les reflets sur les parois d'une tasse éclairée par le soleil peuvent produire au fond une courbe rebroussée, appelée **"caustique"**.
- Mais nous avons négligé jusqu'ici une autre **"grandeur"** caractéristique des classes de rayons lumineux (dans la métaphore archimédienne, elle caractérise la "**torsion**", droite ou gauche, du moment). Grandeur qui appartient en propre à la lumière : c'est la **"constante de Planck réduite"**, que l'on écrit η. Si petite soit-elle, sa prise en compte a une conséquence qualitative importante : la classe des rayons lumineux se découpe en deux **"composantes"** ; la lumière se partage en deux états de **"polarisation circulaire",** "à gauche" et "à droite". Dans les années 1810, Augustin Fresnel a expérimenté la **"biréfringence rotatoire"** des prismes de quartz — qui séparent les deux nappes.

Dès le $\sim$VI$^{\text{è}me}$ siècle, une autre géométrie était apparue, concernant les "figures semblables". Elle est définie par le

Groupe de Thalès

groupe des "**similitudes**" dont le Groupe d'Euclide est "**sous-groupe invariant**".

Pour le Groupe de Thalès, il n'existe qu'une seule classe de couples de points distincts.

Ceci permet de définir le "**rapport de deux distances**", qui appartient nécessairement au "**groupe-quotient Thalès/Euclide**". Groupe que nous décrivons comme "groupe multiplicatif des réels positifs".

Ainsi apparaît l'art de mesurer.

La nature ne permet pas de faire agir le Groupe de Thalès comme le Groupe d'Euclide .

Cependant on rencontre quelques objets qui présentent une régularité "thalésienne" et "non euclidienne". Signe certain de leur origine : elle n'est pas "naturelle", mais "artificielle". Exemple : les coquilles que les bigorneaux savent construire grâce à leur art surnaturel.

Tout aussi artificiellement, nous pouvons faire *agir le Groupe de Thalès* sur les *moments* : il modifie à la fois les forces et les couleurs. De cette action résulte la "proportion" :

$$\hbar \ / \ \text{couleur} = \text{torsion} \ / \ \text{force}$$

d'où les amateurs d'analyse dimensionnelle déduiront que *la couleur est une impulsion.*

Impulsion de quoi ? Des atomes de lumière, disons les "*photons newtoniens*". Et par suite, $\hbar$ sera le *tournoiement* (ou **spin**) commun à tous ces photons.

Nous voyons comment le groupe de Thalès permet de mesurer autre chose que les distances.

Depuis le $18^{ème}$ siècle, cette numérisation est achevée, en ce sens que nous savons construire un groupe de "**matrices**" (tableaux de nombres dont la loi de groupe ressort des "quatre opérations" de l'arithmétique), groupe qui sera déclaré "**isomorphe**" au Groupe d'Euclide.

Les ingrédients qui l'ont permis sont dus à *Pythagore* (son théorème permet de "**numériser l'angle droit"**) et à *Lagrange* (qui a inventé au milieu du $18^{ème}$ siècle les "**coordonnées orthonormales"** que l'on dit *"cartésiennes"*).

Ainsi s'est construite la *géométrie analytique* ; géométrie que Lagrange a complétée en 1788 par la *mécanique analytique.*

3. *Mécanique*, science des machines, devenue *science des mouvements* quand Kepler eut fondé la "**mécanique céleste**" (1609-1621).

Pour la développer, le Groupe d'Euclide sera remplacé par le

Groupe de Galilée

Groupe de Galilée qui agit sur "**l'espace-temps**", constitué des "**événements**" (couples *point-instant*).

Groupe de Galilée qui est lui aussi "**isomorphe à un groupe de matrices**".

Dans la géométrie euclidienne d'Archimède, un moment non nul condamnait les choses au déséquilibre, donc au mouvement.

En géométrie galiléenne, un *moment* sera associé à chaque *mouvement* de chaque chose (doublet latin *momentum—movimentum*).

Moment qui engendre héréditairement un moment de chaque sous-groupe.

Il y a un sous-groupe du Groupe de Galilée que nous subissons et que nous croyons bien connaître : il est constitué des "**translations temporelles**" : "il y a cinq minutes", "dans cinquante ans", etc. Ce groupe, appelons-le

Chronos

Le groupe **Chronos**, nous avons des machines pour en découper des sous-groupes : les *chronomètres*, et autres oscillateurs.

Le *rythme*, élément universel de toutes les musiques, est un choix d'un tel sous-groupe.

C'est aussi **Chronos** qui définit "**l'espace**" : un point *immobile dans l'espace*, c'est une *classe d'événement*s sous l'action de **Chronos** .

Chaque mouvement de chaque chose est pourvu d'un moment galiléen ; donc aussi du sous-groupe **Chronos**. Voilà

l'énergie.

Cette énergie, n'est qu'une partie émergée de l'iceberg *moment*. Pour découvrir les 90% qui restent, il faut envisager d'autres sous-groupes.

C'est ainsi qu'on peut faire apparaître **l'impulsion** et le **tournoiement**, qu'on peut retrouver le "**centre de masse**" sur lequel travaillait déjà Archimède.

Quand deux choses se rencontrent, elles ajoutent leurs moments ; quand elles se séparent, elles le partagent ; elles peuvent en échanger une partie, mais le moment total reste nécessairement constant.

C'est pourquoi nous manipulons le moment avec adresse et prudence quand nous courons entre des obstacles, quand nous prenons ou quittons un escalator, quand nous conduisons une voiture.

Mais un horrible scandale va éclabousser cette belle construction : une analyse élémentaire montre que **Chronos** n'est pas "**invariant**" dans le groupe de Galilée. La géométrie galiléenne ne permet pas de l'y distinguer :

Chronos *, et* **l'espace** *qui va avec, sont donc subjectifs.*

L'espace dans lequel vous vivez, c'est *vous* qui l'avez choisi. C'est celui qui respecte votre confort.

Mais vous acceptez quelquefois de changer de **Chronos**, de quitter l'espace de votre maison pour habiter celui d'une auto, d'un train, d'un avion.

L'expression classique "espace-temps" est donc subjective, elle aussi. Pour respecter la géométrie, il faudra l'éviter : l'ensemble des événements, arène du Groupe de Galilée, le physicien l'appellera

"*Univers*".

De même l'*énergie,* moment du groupe **Chronos**, est aussi subjective que l'espace : l'énergie d'un mouvement que vous faites dans un train, ce n'est pas du tout celle que constatent les observatrices situées dans la campagne environnante.

D'où nous vient donc l'illusion d'un espace absolu ? Le mot "*géométrie*" nous le suggère :

de notre filiation *terrestre.*

Née du tumulte et de la fureur, notre planète-mère a acquis progressivement une *régularité* presque parfaite. Le **Chronos** de tout le monde, n'est que cette régularité de la Terre.

C'est un fait général : une chose isolée qui vieillit tend à acquérir *son propre* ***Chronos***. Un principe premier de la thermodynamique dissipative.

Quelques précisions :
- Le vieillissement de la chose peut se mesurer ; le résultat s'appelle "**entropie**" (Boltzmann, 1877). Le "**second principe de la thermodynamique**" nous affirme que cette entropie ne peut jamais décroître.
- Et le "**premier principe**" ? il est contenu dans un principe dix fois plus riche, la constance du "**moment galiléen**". La croissance de l'entropie doit donc respecter cette décuple constance.
  Les amateurs de "**calcul des variations**" reconnaîtront une circonstance familière : un "**multiplicateur de Lagrange** " $\mathcal{B}$ va intervenir, qui caractérisera chacun des *"états "* où l'entropie ne peut plus croître.
- Géométriquement, $\mathcal{B}$, en dualité avec le moment, habite "**l'algèbre de Lie**" du Groupe de Galilée ; il peut se représenter comme un "**quadri-vecteur**" d'Univers.
- Et le ***Chronos*** *propre de la chose est engendré par* $\mathcal{B}$.

Les astres suffisamment vieux ressemblent à ce portrait, chacun avec son propre $\mathcal{B}$. Mais le **Chronos** associé est souvent un peu plus compliqué que celui que nous avions imaginé tout à l'heure : les planètes et les étoiles *tournent*.

Le fait est là, la Terre tourne, elle aussi.

Le $\mathcal{B}$ de la Terre caractérise sa période de rotation ; le "**jour sidéral**", 23h 56 min. environ.

Cette rotation, la géométrie demande que ce soit celle d'un "solide" — même si la planète conserve une certaine fluidité-plasticité.

La rotation de la Terre casse l'apparente régularité euclidienne, et produit des effets mécaniques paradoxaux : gyroscope de Foucault (1852), alizés et quarantièmes rugissants.

Heureusement, elle ne gêne pas les musiciens.

Pour tout système isolé en équilibre, le quadri-vecteur $\mathcal{B}$ définit aussi la *température d'équilibre* (la mesure de $\mathcal{B}$ est l'inverse de la température absolue).

Même en dehors d'un équilibre thermodynamique achevé, le "vecteur-tempé-rature" $\mathcal{B}$ subsiste localement. Max Planck, en décembre 1900, nous a donné le moyen de le mesurer en observant le rayonnement : c'est ainsi que nous savons qu'autour de notre Galaxie, la température est de 2.73 $^\circ$ Kelvin.

$\mathcal{B}$, c'est enfin la flèche qui nous indique *dans quel sens coule le temps.*

Flèche du temps qui permet le finalisme de la vie : *action sur le futur* grâce au *souvenir du passé.*

Flèche qui interdit toute *action sur le passé* et tout *souvenir du futur* : regrets et espoirs.

4. Le Groupe de Galilée possède une qualité particulière, la "**cohomologie symplectique**", qui se mesure avec une nouvelle quantité particulière :

la masse.

Quelques conséquences de cette cohomologie galiléenne :

- Si la masse n'est pas nulle, chaque mouvement de chaque chose isolée se *décompose* en un couple :
  (**mouvement du centre de masse**, **mouvement autour du centre de masse**).
- On peut construire un nouveau groupe, "**extension centrale**" du Groupe de Galilée ; cette extension, est le *Groupe de Bargmann*.

Chaque "**particule élémentaire**" est associée à une "**classe coadjointe"** du Groupe de Bargmann — qui représente l'ensemble de ses mouvements possibles.

On décrit ainsi les "**points matériels**", les "**particules à spin**", et même les "**photons**" (classés par impulsion, donc par couleur).

L'espace des mouvements de chaque particule est ainsi muni d'une "**géométrie symplectique**", respectée par l'action du Groupe de Galilée.

5. Jusqu'ici, nous avons négligé de prendre en compte

la pesanteur.

Pourtant, nous la pratiquons assidûment depuis l'instant de notre naissance, et nous savons que rien ne lui échappe. Ce caractère universel doit pouvoir s'étayer par la géométrie. Œuvre entreprise en 1916 par Einstein, élaboration de la

Relativité Générale

Dans les années 1920, Einstein parlait de "mollusques de référence" ; Salvador Dali dessinait des "montres molles".

Ces images illustrent l'intervention d'une nouvelle géométrie : la *géométrie différentielle*, qui est définie par le "**groupe des difféomorphismes de classe $\mathbf{C}^\infty$**". Dans le cas des difféomorphismes de l'Univers, nous le nommerons simplement :

Groupe Souple

Comment un tel groupe, beaucoup plus grand que le groupe de Galilée (on dit que "*sa dimension est infinie*"), peut-il intervenir dans la physique ?

Très simplement, si on veut bien examiner la simple "chute des corps".

Comme nous l'a enseigné Galilée, tout "**point matériel**" qui tombe emprunte "**l'accélération de la pesanteur**". Cette loi s'applique aussi bien aux astres qu'à la pomme de Newton. Or on peut l'écrire avec la nouvelle géométrie :

- La *pesanteur* peuts'interpréter comme objet souple : une "**connexion affine symétrique**" **D**.
- La "**ligne d'Univers**" d'un point matériel, associée à sa "**masse**", devient aussi un *objet souple* **T** : un "**2-tenseur-distribution-contravariant-symétrique**".
- Alors la "**loi de la chute des corps**" devient :
  " **TDZ = 0 si Z est un 1-tenseur $\mathbf{C}^\infty$ à support compact** " ;
  en abrégé :

$$\mathbf{TD = 0}.$$

L'intérêt de cette écriture, c'est qu'elle peut devenir une propriété universelle de la matière :

- $\mathbf{TD = 0}$ implique d'abord la "**conservation de la masse**" au cours du mouvement.
- Si deux particules $\mathbf{T}_1$ et $\mathbf{T}_2$ tombent indépendamment, leur chute commune sera décrite par la distribution $\mathbf{T} = \mathbf{T}_1 + \mathbf{T}_2$, nouvelle solution de $\mathbf{TD = 0}$.
- Si ces particules se choquent, l'équation $\mathbf{TD = 0}$ implique aussi toutes les "l**ois des collisions**", celles qui ont été formulées au $17^{\text{ème}}$ siècle (Galilée, Mariotte, Huygens).
- Un seul objet $\mathbf{T}$ **,** solution de l'équation $\mathbf{TD = 0}$, peut ainsi peupler l'Univers d'une *infinité* de particules qui tombent en s'entrechoquant : *vide, atomes* et *clinamen* selon Démocrite et Lucrèce.
- Une *loi de probabilité* sur les solutions du type précédent permet de définir une solution "moyenne" $\mathbf{T}$, géométrisant la"**théorie cinétique des gaz**". Apparition de la "**pression**".
- Ces solutions, et d'autres, permettent de modéliser les divers "**milieux continus**". Dans ces cas-là, $\mathbf{T}$ peut s'exprimer avec la *densité*, la *vitesse* et la *contrainte* du milieu. Alors $\mathbf{TD = 0}$ produit le système des *équations d'Euler* (1755) ; équations écrites ici "**linéairement**".

Par ailleurs, un simple calcul montre que *toute solution* isolée de l'équation $\mathbf{TD = 0}$, associée à une solution tensorielle $\mathbf{Z}$ de l'équation $\mathbf{DZ = 0,}$ définit une "***grandeur conservée***".

Si l'on se trouve dans une station spatiale, en "**impesanteur**", il y a beaucoup de solutions de l'équation $\mathbf{DZ = 0,}$ donc beaucoup de grandeurs conservées pour chaque mouvement. On connaît déjà la *masse* : on trouve en plus *l'impulsion*, le *tournoiement* , le *centre de masse*. Dix grandeurs au total sur lesquelles peuvent s'appuyer les astronautes dans leurs évolutions.

L'objet protéiforme $\mathbf{T}$ que nous venons de rencontrer, nous l'appellerons :

dynamique de la matière

et l'équation $\mathbf{TD = 0}$ méritera le nom de "**loi de la dynamique**".

Le savoir-faire des mécaniciens, c'est de joindre à cette "l**oi de la dynamique**" des **"lois de comportement"** concernant les matériaux qu'ils étudient.

Lois qu'ils choisissent invariantes par l'action du Groupe de Galilée — sous-groupe du Groupe Souple. Ont-ils raison ? Nous verrons.

Ainsi peuvent se construire la "**théorie de l'élasticité** ", la "**mécanique des fluides**". L'usage des distributions permet aussi de décrire les "**coques**" de la chaudronnerie, la "**tension superficielle**" qui permet à des insectes de marcher sur l'eau ; la simple "t**héorie des cordes**", tendues ou vibrantes ; de définir les "**forces**" qu'elles transmettent, de "**composer**" ces forces. Pensons à une toile d'araignée agitée par l'impact d'une mouche.

Et ça fonctionne parfaitement dans le cas des "**phénomènes dissipatifs**", qui produisent des résultats irréversibles.

6. Mais un nouveau problème apparaît : pas question de traiter ainsi la loi de l'équivalence "**énergie-chaleur**" — simplement parce que *l'énergie a disparu* !

Pour la retrouver, il faudra *sacrifier le Groupe de Galilée.*

Revenons aux "particules élémentaires" que nous avions déduites de ce groupe.

Il y avait une difficulté concernant les *photons* : leur vitesse était rigoureusement infinie. Or, dès 1676, Rœmer avait mesuré la vitesse c de la lumière et constaté qu'elle était finie — si grande soit-elle. Et la même pour toutes les couleurs.

Pour avoir de "bons photons", une seule solution : trouver un substitut du groupe de Galilée qui implique cette valeur c. C'est possible, et on peut trouver ainsi le "**groupe de Poincaré**", sous-groupe du Groupe Souple. Sa dimension est 10, comme celle du Groupe de Galilée ; sa géométrie, la

Relativité Restreinte

fournit de nouvelles particules, tout aussi symplectiques, mais plus conformes à l'expérience ; les photons vont maintenant à la bonne vitesse c.

Relativité Restreinte qui apparaît comme épistémologiquement postérieure à la Relativité Générale — même si elle est historiquement antérieure.

Cette géométrie est légale depuis 1983 : la vitesse c de la lumière a été ajustée arbitrairement (un nombre entier de mètres par seconde) ; le mètre-étalon a donc été relégué au musée.

Voilà enfin un événement bien daté de l'histoire de la géométrie au $20^{\text{ième}}$ siècle.

Une nouveauté de la "Relativité Restreinte" : *la masse d'une chose dépend de son mouvement* — selon la formule $\mathbf{m} = \mathbf{E}/c^2$, $\mathbf{E}$ désignant l'énergie retrouvée.

Résultat établi à la fin du $19^{\text{ème}}$ siècle par Henri Poincaré dans le cas de la lumière ; en 1905, Albert Einstein a déclaré que cette formule était générale, et il a eu la bonne idée d'y chasser le dénominateur.

Mais comment se fait-il que nous ayons une grandeur conservée en moins, masse et énergie ayant perdu leur indépendance ?

Parce que le Groupe de Bargmann permettait d'utiliser subrepticement la "**masse atomique**" des chimistes, dite aussi "**quantité de matière**". Quantité qui varie peu dans les conditions généralement réalisées sur Terre.

La géométrie de Galilée-Bargmann était celle du paradis écologiste : pas de radioactivité, pas d'énergie nucléaire. Mais la géométrie de Poincaré respecte mieux la nature *"telle qu'elle est"*.

Quelques conséquences de la Relativité Restreinte :

- Les photons n'ont plus de *couleur* propre, leur couleur est relative à qui les regarde (effet Doppler-Fizeau).

- La "**cohomologie symplectique**" a disparu : disparition de la **décomposition barycentrique**. Le *solide mobile*, dont Euler et Poinsot avaient calculé les mouvements libres "autour du centre de masse", n'existe plus : pour décrire les mouvements des objets les plus durs, il devient nécessaire de prendre en compte leurs propriétés **élastiques** ou **plastiques**.
- Mais la loi de la dynamique $\mathbf{TD = 0}$ reste inchangée. Les mécaniciens pourraient tout aussi bien utiliser le Groupe de Poincaré pour écrire les lois de comportement, et ils n'auraient plus de problème avec l'énergie, puisqu'ils pourraient admettre que $\mathbf{E}= \mathbf{m}c^2$.

Mais la place du Groupe de Poincaré dans le Groupe Souple pose problème.

Pourrait-on le définir comme "régularité de l'impesanteur" ? Eh bien non, ce qu'on obtient ainsi, c'est un groupe plus grand, dont la dimension est 20 : le "**groupe affine**".

Voici : le Groupe de Poincaré est bien la régularité d'un objet souple $\mathbf{g}$, objet d'un type nouveau : une "**métrique hyperbolique normale**".

La vitesse c de la lumière est inscrite dans $\mathbf{g}$, ce qui explique comment c peut se transmettre au Groupe de Poincaré.

métrique d'Univers

Examinons donc cet objet fondamental de la physique — dont l'étude a commencé avec les travaux d'Hermann Minkowski (1908).

- Comme son nom l'indique, la *métrique* $\mathbf{g}$ permet de *mesurer* les distances et les durées, selon une méthode due à Bernhard Riemann (1853).
- La thermodynamique des "vieilles choses" telles que la Terre implique toujours un "vecteur-température" $\mathcal{B}$, qui appartient maintenant à "**l'algèbre de Lie du Groupe de Poincaré".** Dans le cas des choses "en cours de vieillissement", $\mathcal{B}$, devenu objet souple, engendre avec $\mathbf{g}$ un nouvel objet $\mathcal{B}\mathbf{g}$ ("**la dérivée de Lie de la métrique g par le vecteur** $\mathcal{B}$"). Si $\mathcal{B}\mathbf{g}$ n'est pas nul, la "**dissipation"** va apparaître.
  Ainsi peuvent se géométriser la *conduction de la chaleur* (équations de Fourier, 1822), la *théorie de la viscosité* (équations de Navier, 1822) ; etc.
- La métrique définit enfin la *pesanteur*, comme la seule connexion$\mathbf{D}$ qui vérifie l'équation $\mathbf{Dg = 0}$.

Mais la pesanteur ainsi obtenue ne correspond à l'expérience que dans une région *pas trop grande* : elle exigerait que la Terre soit *plate* — or nous savons qu'elle est ne l'est pas.. . .

Le Groupe de Poincaré et sa Relativité Restreinte ne sont donc qu'une *approximation microscopique.*

7. A grande échelle, comment utiliser le Groupe Souple ?

Demandons-lui ce que nous demandions jusqu'ici aux groupes d'Euclide, Galilée ou Poincaré : agir sur le réel sans modifier la physique. Autrement dit :

agir sur le possible.

Choisissez un point quelconque de l'Univers. Vous trouverez peut-être un élément du Groupe Souple qui vous y emmènera. Si vous le faites agir, vous aurez fait un grand voyage, et même un voyage dans le temps.

Comment ce voyage est-il *possible* ? Simplement parce que *vous ne vous apercevrez de rien.* Vous aurez été accompagné par la métrique, par la pesanteur ; par votre environnement ; par le reste du monde.

L'action du Groupe Souple transforme le possible en possible parce qu'*elle ne change rien,* parce qu'elle est *inobservable*...

... nouvelle épistémologie de la géométrie.

Soyons plus précis : envisageons l'ensemble $\mathcal{G}$ des *métriques possibles* de l'Univers. Chaque classe de $\mathcal{G}$ sous l'action du Groupe Souple, appelons-la

un cosmos

Physis

En souvenir d'Empédocle, appelons

L'ensemble Φ *de tous les Cosmos.* Notre Cosmos à nous, ce n'est plus que *l'un des points de la Physis.*

Nous pouvons munir $\mathcal{G}$ de la "**différologie compacte**", qui est invariante par l'action du Groupe Souple. Elle descend sur la Physis Φ , en la munissant de la "**différologie quotient**". Grâce à cette construction, la mécanique empirique que nous venons de rencontrer :

- la dynamique de la matière **T** est une distribution qui vérifie $\mathbf{TD = 0}$ **, D** *désignant la connexion qui dérive de* **g** *par la condition* $\mathbf{Dg = 0}$

se traduit géométriquement :

- **T** est un tenseur de Φ.

Ainsi la dynamique **T** *habite la* Physis Φ ; elle est définie au point φ_o qui est "notre Cosmos" ; c'est son "**relevé**" sur l'espace $\mathcal{G}$ des métriques que nous expérimentons.

Mais pourquoi notre Cosmos serait-il le seul à bénéficier d'une dynamique ? Pour être *"scientifiquement correct"*, envisageons que **T** soit défini *sur toute la Physis.*

Interprétation : le *possible* doit être astreint à une *loi* qui exprime la *dynamique* à partir de la *métrique.*

C'est avec une idée analogue que Cavendish, expérimentant en 1798 sur la pesanteur, a réussi à mesurer la masse de la Terre. Les mesures actuelles ne diffèrent des siennes que de 1%.

La loi qu'il éprouvait, c'était "*l'attraction universelle*", celle que Newton avait déjà induite des *lois de Kepler.*

Problème : peut-on réécrire la *"loi de l'attraction"*, sans rien perdre de ses implications keplériennes et cavendishiennes, pour prolonger le tenseur **T** sur la Physis ?

Oui ; une solution, c'est *l'équation d'Einstein* (1915-1917). Elle munit la Physis d'une "**1-forme fermée**" ; sur l'Univers, elle s'écrit $\mathbf{T = S(g)}$**, S** étant un "**opérateur différentie**l".

Elle se vérifie avec une précision extrême. Des variantes sont possibles, mais aucune n'a franchi le seuil expérimental.

La loi d'Einstein contient deux "**constantes universelles**" : la *constante de Newton* et la *constante cosmologique ;* la valeur de leur couple est mesurée depuis 1999 à partir des observations de supernovæ lointaines.

Mais le géomètre est parfaitement indifférent à la valeur de ces constantes, parce qu'il peut attribuer à chacune une valeur arbitraire, 1 par exemple. Opération qui revient à choisir des "**unités gravitationnelles**" pour la *masse* et pour les *durées-distances.*

Unités très grandes : elles donnent par exemple à la constante de Planck l'ordre de grandeur $\hbar \approx 10^{-120}$.

8. Plaçons-nous maintenant à cette échelle des unités gravitationnelles, et jetons un coup d'œil sur les

Modèles cosmologiques

Si l'on examine diverses photographies du ciel où ne sont portés que les objets suffisamment lointains, on constate que les galaxies présentent une "densité" qui varie peu avec la direction.

Autrement dit : les galaxies lointaines peuvent être décrites par une "statistique" possédant la régularité de la sphère céleste. Groupe isomorphe à **O(3)** ; "**isotropie**", dit-on.

Mais nous ne testons ainsi que ce qui est visible depuis la Terre : on peut présumer que l'isotropie n'est qu'un sous-groupe d'une "*régularité cosmologique*" **C**.

Une hypothèse concernant ce groupe **C** : il pourrait être isomorphe à **O(4)**. Sous-groupe "**compact**" du Groupe Souple, qui permettrait alors de construire la "**moyenne**" **T** de la dynamique réelle.

Ce **T** moyen peut se décrire comme un "fluide à pression nulle", schéma d'une "poussière de galaxies". L'équation d'Einstein permet d'associer **T** à une métrique **g** qui possède aussi la régularité **C**.

Ainsi produit-on un *modèle cosmologique*, proposé dès 1922 par Alexandre Friedmann.

L'action du groupe **C** *classe* les événements. En compréhension, chaque classe est une *"date cosmique"* ; en extension, c'est *"l'espace"* à cette date-là ; espace muni de la géométrie induite de **C** — géométrie *éternelle*, donc. "Espace co-mobile", disent les astronomes.

Cette géométrie est celle d'une hypersphère **S3** dont le rayon **R** croît avec la date : "**expansion de l'Univers**". À la date "**présente**", l'espace est "**courbe**", mais pas beaucoup, parce que **R** est grand.

Depuis 1929, ce modèle se contrôle par l'observation d'un "**décalage spectral**" des galaxies lointaines, fonction de leur "**distance**" (Edwin Hubble disait imprudemment "proportionnel à la distance").

Souvenons-nous maintenant d'un théorème grâce auquel Archimède a pu calculer l'aire de la sphère (il concerne l'aire des "**calottes sphériques**").

On peut l'étendre en dimension quelconque ; il en résulte que la répartition tridimensionnelle des astres, supposée *équipartie* sur l'hypersphère **S3** , se projettera **équipartie** sur tout disque équatorial **D2**. Ceci permet de calculer le volume de l'espace ; et aussi de contrôler visuellement les aspects *non-géocentriques* de la régularité **C**.

On peut perfectionner le modèle de Friedmann en tenant compte du *rayonnement cosmologique* (observé depuis 1964) ; rayonnement caractérisé par un vecteur-température β pointant vers le futur ; rayonnement dont la dynamique comporte *densité* et *pression* de radiation.

Bien entendu β bénéficie de la régularité **C** ; sa mesure $|\beta|$ (la "**température réciproque**"), est proportionnelle au rayon **R** de l'espace : refroidissement perpétuel ; 2.73 $^\circ$K aujourd'hui.

La dérivée de Lie β **g** n'est pas nulle, ce qui laisse présager une production d'entropie (le "flux d'entropie" est un moment du Groupe Souple) : "**interactions dissipatives**" donc entre la matière et le rayonnement.

9. Revenons à notre planète, qui nous a déjà fourni quelques enseignements : nous avions remarqué que *la Terre tourne*, puis qu'elle *n'est pas plate* ; nous pouvons préciser maintenant qu'*elle est ronde*. Cette rondeur, c'est une régularité de la Terre.

Compte tenu de son aplatissement, il s'agit d'une régularité de type "**Chronos** x **O(2)**" Les écarts de la pesanteur à cette régularité n'ont pu être détectées qu'à la fin du $18^{ème}$ siècle : expériences de Bouguer et Maskelyne au voisinage des monts Chimborazo et Shiehallion.

Mais la Terre est déjà "très sphérique", la régularité "**Chronos** x **O(3)**" est en cours d'apparition.

La simple géographie nous donne donc des exemples de régularités *plus grandes que* ***Chronos*** qui apparaissent spontanément. Voici des exemples à d'autres échelles.

- Une régularité euclidienne de type **A5** (groupe à 60 éléments) appartient à deux des polyèdres de Platon, et aussi aux ballons de football. Or cette régularité artificielle appartient aussi à des objets produits spontanément par la nature : dans la flamme d'une bougie qui fume apparaît une molécule possédant cette régularité-là : le carbone $C60$, découvert en 1985, qu'on appelle aussi "footballène".
- Depuis 4.7 milliards d'années, des phénomènes dissipatifs se produisent dans le système solaire (collisions, marées) ; une régularité de type "Fibonacci" est apparue : les *périodes* des planètes s'éloignent peu d'une "**progression géométrique**" dont la raison serait le "**Nombre d'or**". Même résultat pour les systèmes de satellites de Jupiter et de Saturne.

C'est parce que le Nombre d'or est "**algébrique de degré 2**" qu'il caractérise une régularité de type $\mathbf{T}^2$, correspondant à une diminution de la dissipation par les marées planétaires ou solaires.

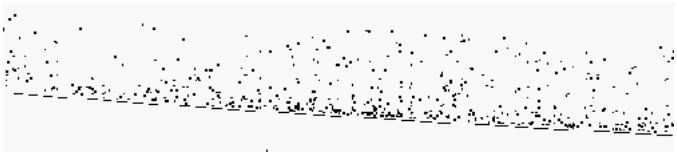
Exemples divers de régularités apparues "naturellement" — ayant la vertu de "modérer" la dissipation.

Apparues comment ? La thermodynamique est encore incapable de nous l'expliquer avec précision : elle propose simplement l'expression : "**structures dissipatives**".

La régularité cosmologique **C** pourrait donc, elle aussi, avoir été acquise à la suite des processus dissipatifs imposés par le modèle même.

Dans ce cas, son extrapolation stricte vers le passé ne serait pas légitime. L'isotropie apparente du ciel lointain ne serait qu'une *anisotropie en voie de disparition.*

Anisotropie dont on peut relever une trace : la tendance des quasars les plus proches à se grouper en *"hyper-amas"* aplatis et parallèles — verticaux sur la projection archimédienne suivante :



Tendance dont l'optimisation permet d'évaluer les paramètres cosmologiques — en agrément avec divers tests cosmologiques récents.

Elle ne fait pas disparaître la régularité cosmologique **C** , mais la réduit simplement à une régularité **non-géocentrique** de type **O(3)**, qui pourrait avoir été primordiale.

10. Jusqu'ici, nous avons négligé les phénomènes magnétiques et électriques. Mais on peut les prendre en compte avec un simple ajustement de la géométrie : on remplacera le Groupe Souple par le

Groupe Électro-Souple,

dont voici la construction :

- L'*Électro-Univers* sera le produit direct de l'Univers **U** par un cercle ; c'est une variété $\underline{\mathbf{U}}$ de dimension 5 (Theodor Kaluza, Oskar Klein, 1919-1926).
- En faisant tourner simultanément tous les cercles du même angle, on produit le *Groupe Électrique.*
- Le *Groupe Électro-Souple* sera le "**normalisateur**" du Groupe Électrique dans le "**groupe des difféomorphismes**" de $\underline{\mathbf{U}}$.

Le Groupe Electro-Souple est une "**extension**" du Groupe Souple : les objets souples sont donc automatiquement *électro-souples* ; la nouvelle géométrie est *compatible* avec la précédente.

Le "**groupe quotient**" contient les "**transformations de jauge de seconde espèce**" de l'Électrodynamique, et aussi les "**conjugaisons de charge**". La géométrie Electro-Souple est ainsi apte à décrire simultanément *matière* et *antimatière.*

Cette géométrie recourt à une *Electro-Métrique* $\underline{\mathbf{g}}$, métrique de *l'Electro-Univers* $\underline{\mathbf{U}}$, qui peut être choisie :

- invariante par le Groupe Électrique ;
- donnant à tous les cercles la même longueur.

On pourra alors définir *l'Electro-Physis* correspondante, et *l'Électro-Dynamique* $\underline{\mathbf{T}}$ comme tenseur de l'Electro-Physis.

Voici quelques conséquences de ces hypothèses.

L'Electro-métrique $\underline{\mathbf{g}}$ s'exprime en complétant la métrique souple $\mathbf{g}$ par un objet électro-souple $\mathbf{A}$, le "**potentiel électromagnétique**" :

$$\underline{\mathbf{g}} = (\ \mathbf{g}\ ,\ \mathbf{A}\ )$$

De cette Électro-Métrique, on dérive l'*électro-pesanteur* $\underline{\mathbf{D}}$, qui complète la pesanteur $\mathbf{D}$ par le **champ électromagnétiqueF** :

$$\underline{\mathbf{D}} = (\ \mathbf{D}\ ,\ \mathbf{F}\ ).$$

$\mathbf{F}$, objet souple, est automatiquement solution des **équations de Maxwell homogènes**.

L'*Électro-Dynamique* $\underline{\mathbf{T}}$ complète la dynamique $\mathbf{T}$ par un "**vecteur-distribution**" $\mathbf{J}$, le **courant électrique** :

$$\underline{\mathbf{T}} = (\ \mathbf{T}\ ,\ \mathbf{J}\ ).$$

Alors la *loi de l'Électro-Dynamique* $\underline{\mathbf{T}}\ \underline{\mathbf{D}} = \mathbf{0}$ se coupe en deux :

- Les **forces électromagnétiques** interviennent dans l'équation de la dynamique qui devient :
$$\mathbf{TD} + \mathbf{JF} = \mathbf{0}\,;$$
- le courant électrique $\mathbf{J}$ vérifie l'équation
$$\mathbf{JD} = \mathbf{0}$$
qui s'interprète comme *conservation de l'électricité*. La **charge électrique** ainsi conservée est un *moment du Groupe Électrique*.

Enfin l'Électro-Dynamique se relève de l'Electro-Physis par "l'équation d'Einstein pentadimensionnelle" $\underline{\mathbf{T}} = \underline{\mathbf{S}}(\underline{\mathbf{g}})$:

- dans l'écriture de l'équation d'Einstein**,** la dynamique souple $\mathbf{T}$ doit être augmentée de la "**densité tensorielle de Maxwell-Poynting**" ;
- $\mathbf{J}$ est donné par les "**équations de Maxwell non homogènes**".

Ce système $\underline{\mathbf{T}} = \underline{\mathbf{S}}(\underline{\mathbf{g}})$ entraîne bien entendu $\underline{\mathbf{T}}\ \underline{\mathbf{D}} = \mathbf{0}$.

Cette Électrodynamique géométrique correspond parfaitement à l'expérience et à la technologie : équivalence des aimants et des courants (André Marie Ampère, 1827), générateurs électriques (Zénobe Gramme, 1865), utilisables réciproquement comme moteurs (1873) , émission et réception des ondes radio (1894), etc.

C'est d'ailleurs l'analyse géométrique de l'électrodynamique classique qui a permis à Henri Poincaré de découvrir en 1905 le groupe qui produit la Relativité Restreinte.

Le fait que l'attraction gravitationnelle ne soit pas supplantée par la répulsion électrique est une limite drastique de la **charge électrique** de l'espace : on admet généralement que cette charge universelle est *nulle*. Plusieurs scénarios peuvent être évoqués pour géométriser cette hypothèse.

11. Définissons maintenant un nouvel Universel géométrique : les

états d'un groupe

Pour chaque groupe **G** , un *état* est une fonction μ définie sur **G** , à valeurs complexes, prenant la valeur 1 au point neutre, telle que les $\mu(\mathbf{g}_j^{-1}\mathbf{g}_k)$ soient les éléments d'une "**matrice positive**" — quel que soient les $\mathbf{g}_j$ choisis dans **G**.

Voici quelques-unes de leurs propriétés :
- Le conjugué d'un état, le produit de deux états sont des états.
- Les états de **G** constituent un "convexe compact" — engendré par les "états purs" : ceux qui ne sont pas le milieu de deux autres.
- Chaque état définit une "**représentation unitaire**" du groupe sur un "**espace de Hilbert**".
- Réciproquement, tout espace de Hilbert, toute représentation unitaire d'un groupe peuvent se reconstruire à partir d'un état.

Les "**représentations irréductibles**" sont celles qui sont associées à un "**état pur**".

12.

calcul des probabilités

A quoi peuvent servir tous ces états ? D'abord à fonder le Pour "probabiliser" un ensemble **X ,** il suffit de choisir un "**groupe-test**" : un groupe multiplicatif **A** de fonctions α à valeurs complexes de module 1, définies sur **X** .

Fermons **A** pour les "*combinaisons linéaires*" et "*limites uniformes*" : nous obtenons l'ensemble des "**fonctions-test**" de **A**. Appelons
- "**certitudes**" les fonctions sur **A** associées à chaque point **x** de **X** par la correspondance α—> $\alpha(\mathbf{x})$;
- "**hasards**" les fonctions du "**convexe fermé**" engendré par les certitudes (fermé pour la "**convergence uniforme sur les parties finies**").

Tous ces hasards sont des états de **A**. Tout hasard **p** possède un seul prolongement linéaire continu à l'ensemble des fonctions-test αg le résultat $\mathbf{p}(\alpha)$ sera l' "**espérance mathématique**" ou "**valeur moyenne**" de α dans le hasard **p**.

Voici des exemples :
- Si **X** est fini, tout hasard sur **X** est donné par une formule
  $\mathbf{p}(\alpha) = \mathbf{p}_1\ \alpha(\mathbf{x}_1) + \ldots + \mathbf{p}_n\ \alpha(\mathbf{x}_n)$,

  les $\mathbf{p}_j$ étants positifs, de somme 1 : la notion de hasard implique celle de "**probabilité**".
- Prenons un exemple infini : **X**, ce sera $\mathbf{R}^n$; on choisit comme "**groupe-test"** l'ensemble des fonctions $\alpha_y : \mathbf{x} \text{ —> } \mathbf{e}^{iyx}$, **y** étant arbitraire dans le dual de $\mathbf{R}^n$.
  Alors les "**hasards**" contiennent les "**lois de probabilité**" classiques (représentées par leurs "**fonctions caractéristiques**" au sens de Liapounov-Lévy).
  Un point intéressant : il existe d'autres hasards probabilisés par ce groupe-test, comme celui qui est *équiparti* sur $\mathbf{R}^n$ ou sur un sous-réseau arbitraire. Ils interviennent dans les phénomènes quantiques — même s'ils sont *incompatibles avec la théorie de la mesure.*

Soient **X** , **X'** deux espaces probabilisés par des groupes-tests **A** , **A'**. Une application **F** de **X** dans **X'** sera dite "**propre**" si l'image réciproque par **F** d'un élément de **A**' est fonction-test sur **X**.

Alors l'image réciproque d'une fonction-test de **X'** par une application propre est fonction-test de **X** ; l'image directe d'un hasard de **X** (définie par dualité) est un hasard de **X'**.

Ces "**images de hasards**" sont utilisées dans les "**méthodes de Monte-Carlo**". Elles permettent à un groupe qui agit "**proprement"** sur **X** d'inclure les *hasards* dans sa *géométrie*.

13. D'autre part ces états de groupe permettent d'accéder à la procédure de

quantification géométrique

qui ne sera qu'esquissée ici.

Un objet quantique isolé sera caractérisé par :
- un groupe **S** — **la source**
- une **classe coadjointe X** de la source **S** (qui s'interprétera comme *"espace des mouvements classiques"*)
- un **fibré principal en cercles** Ξ au dessus de **X** muni d'une **structure de contact** ϖ d'où dérive la **forme symplectique** ω de **X**. La **circulation** de ϖ su les cercles prend la valeur constante $2\pi\eta$; l'action de **S** sur **X** se relève sur Ξ en préservant ϖ — ce qui crée une "**extension centrale**" Σ de **S**.
- si l'objet est isolé, un morphisme du groupe de Galilée (resp. Poincaré) dans **S**

Appelons "**observation**" tout sous-groupe **A** de **S** engendré par une **sous-algèbre de Lie abélienne**. Chaque observation pourra s'interpréter de deux façons :
- par son action sur Ξ induite de celle de Σ ;
- par une autre action sur Ξ, déterminée par ϖ, et conservant individuellement les fibres ; elle produit un "**groupe-test**" de **X**.

Si μ est un **état** de la source et **A** une observation, μ induit sur **A** un état du états quantiques groupe-test associé. Les

états quantiques

seront des états de **S**, ayant la propriété d'induire un "**hasard**" pour chaque observation **A**. Hasard qui s'interprétera comme "**spectre**" du moment de **A**.

Ces constructions permettent de produire les principaux objets de la mécanique quantique :
- *Équations d'onde* (Schr$\acute{}$odinger, Pauli, Klein-Gordon, Dirac, Maxwell), *engendrant chacune une solution de* $\mathbf{TD = 0}$
- *Espaces de Hilbert, vecteurs d'état, opérateurs-densité*
- *opérateurs self-adjoints*
- *Produit tensoriels hilbertiens pour les systèmes*
- *bosons* et *fermions*, *condensations de Bose-Einstein*, *non-séparabilité des états de fermions.*

Bibliographie

[1] J. M. Souriau et al., *Géométrie symplectique différentielle. Applications* - Coll. Internat. CNRS "Géométrie différentielle", Strasbourg, p.53, 1953.

[2] J. M. Souriau et al., *Géométrie et relativité* - Ed. Hermann, coll. "Enseignement des Sciences", 1953.

[3] J. M. Souriau et al., *Définition covariante des équilibres thermodynamiques* - Suppl. Nuov. Cimento, 1, 4, pp.203-216, 1966.

[4] J. M. Souriau et al., *Modèles classiques quantifiables pour les particules élémentaires* - C.R. Acad. Sc. 263, p.1191, 1966.

[5] J. M. Souriau et al., *Quantification géométrique* - Comm. Math. Phys. 1, pp.374-398, 1966.

[6] J. M. Souriau et al., *Quantification géométrique. Applications* - Ann. Inst. H. Poincaré VI, 4, p.311-341, 1967.

[7] J. M. Souriau et al., *Structure des systèmes dynamiques* - Ed. Dunod, Coll. "Dunod Université", 1969.

[8] J.M. Souriau et al., *Variétés symplectiques et cohomologie en mécanique* - Rencontres Math. Phys., Dep^t Math. Univ. Lyon-Villeurbanne, 8, Suppl., 1971.

[9] J. M. Souriau et al., *Construction explicite de l'indice de Maslov. Applications* - "4th Internat. Colloquium on Group Theoretical Methods in Physics", Nijmegen, 1975 - Springer Lect. Notes in Physics 50, p.117-148, 1976.

[10] J. M. Souriau et al., *Faut-il prendre au sérieux la constante cosmologique ?* - Journ. Relativistes, Bruxelles 1976, Publ. Univ. libre de Bruxelles, p.215-229, 1977.

[11] J. M. Souriau et al., *Thermodynamique et géométrie* - "Differential Geometrical Methods in Mathematical Physics II", Bonn - Springer Lect. Notes in Math. 676, 30 pages, 1977.

[12] J. M. Souriau et al., *Groupes différentiels* - Springer Lect. Notes in Math., 836, p.91-128, 1980.

[13] J. M. Souriau et al., *A Possible Large-Scale Anisotropy of the Universe* (avec *H.H. Fliche* et *R.Triay*) - Astron. and Astrophysics 108, p.256-264, 1982.

[14] J. M. Souriau et al., *Heat, Cold and Geometry* (avec *P. Iglesias*) - "Differential Geometry and Mathematical Physics", M.Cahen ed^r, p.37-68, Reidel Pub. C$^\circ$, 1983.

[15] J. M. Souriau et al., *Groupes différentiels et Physique mathématique* - "Feuilletages et quantification géométrique", Journées Internat. Soc. Math. de France 1983, Lyon - Coll. Travaux en cours, II, p.73-119, Ed. Hermann, 1984.

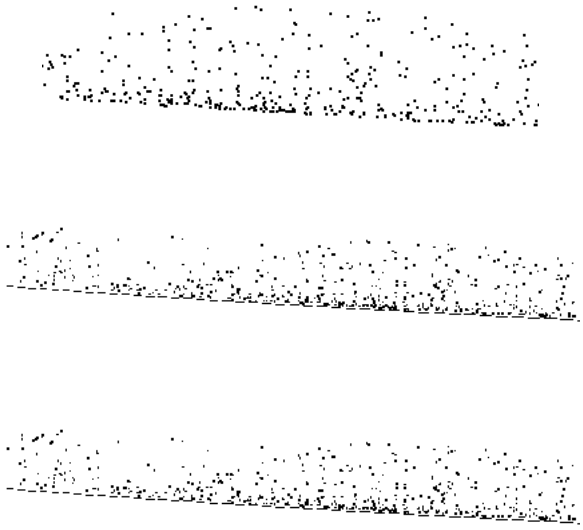
[16] J. M. Souriau et al., *La structure symplectique de la mécanique décrite par Lagrange en 1811* - "Journées d'Histoires des Sciences", Vieille Charité, Marseille,1985 - "Mathématiques et Sciences humaines", 24ème année, n ° 94, pp.45-54, 1986.

[17] J. M. Souriau et al., *Milieux continus de dimension 1, 2 ou 3 : statique et dynamique* -13ème Congrès Français de Mécanique, Poitiers Futuroscope, 1997.

[18] J. M. Souriau et al., *Structure of Dynamical Systems, a symplectic View or Physics* - Birkhäuser, Progress in Mathematics, 1998.

[19] J. M. Souriau et al., *Grammaire de la Nature - à paraître*, 2002.

remords



Géométrisation de la théorie physique : sur la genèse d'un problème

Alain Michel

Le problème consistant à " géométriser " une question, une situation, un domaine, donnés ne reçoit de sens que si l'on suppose une référence, au moins implicite, au champ physique. Il ne peut recevoir de réponse tant qu'on n'a pas éclairci la question de l'espace. Aussi son traitement s'inscrit-il naturellement dans les tentatives déployées par les mathématiciens pour parvenir à un concept satisfaisant de l'espace. C'est un pan de l'histoire de cette mise en forme connexe de la géométrisation et de l'espace qu'on voudrait restituer. La difficulté est de le faire fidèlement. On peut tenir pour établi un double fait.

Le premier est que les problèmes qui sont à l'origine des importantes recherches qui se développent dans le champ de la géométrie, et surtout de la géométrie différentielle, après Riemann, ont leur source dans un problème plus général. Celui-ci n'est pas seulement, ou strictement, mathématique, mais un problème qu'on pourrait qualifier, avec les précautions d'usage, de " philosophique ", et que les auteurs ont le plus ordinairement appelé " problème de l'espace ".

Cela est vrai même d'un point de vue historique. C'est un second fait que les mathématiciens et physiciens théoriciens qui, en cette première moitié du XX$^{\text{ème}}$ siècle, s'occupent de ce problème, créant de nouveaux outils mathématiques et reconstruisant la théorie physique, sont parfaitement conscients de la nature de l'enjeu. Tous inscrivent leurs efforts dans une réflexion sur l'espace.

Dans la deuxième moitié du XIX$^{\text{ème}}$ siècle, et durant tout le début du XX$^{\text{ème}}$ siècle, philosophes et mathématiciens traitent volontiers du " problème de l'espace ". Contentons-nous de citer les deux cas, mais exemplaires, d'Hermann Weyl et de Rudolf Carnap [1][2], tous les deux auteurs, sous ce titre, d'études également décisives pour la suite.

On s'accorde généralement à reconnaître que le problème en lui-même est de nature philosophique, mais, depuis Kant, et surtout à partir de Riemann, ce sont les mathématiciens qui ont contribué à en fixer et à en préciser les termes. C'est un trait de notre temps que le développement prodigieux des mathématiques à la fois permette un traitement toujours plus raffiné et plus systématique des termes du problème et recouvre, jusqu'à l'oubli, sa signifi-

cation fondamentale —sa signification de problème. Il faut sans doute le regretter, et savoir reconnaître que la réflexion doit se mouvoir ici entre deux pôles irréductibles. L'impossibilité de réduire l'un à l'autre est justement ce qui donne au problème sa profondeur, comme l'ont compris les deux auteurs cités.

Ils se seraient probablement accordés aussi pour reconnaître les origines kantiennes de ce qu'ils nomment le " problème de l'espace " – son caractère de " problème transcendantal ". On sait le rôle de la physique, surtout de la mécanique, dans la réflexion kantienne, et c'est bien dans le cadre de la physique, non de la mathématique, que l'espace reçoit, dans la doctrine de l'idéalisme transcendantal, son statut de problème.

Il s'agira ici d'un essai de mise en perspective philosophique des efforts déployés par les mathématiciens pour parvenir à un traitement du " problème de l'espace ". C'est dans cette perspective d'une actualisation de la réflexion transcendantale qu'on proposera quelques réflexions à propos du grand travail de généralisation qui intervient au début du XX$^{\text{ème}}$ siècle, en relation avec la théorie de la relativité. A bien des égards, ce problème émigre, dans toute cette première moitié du XX$^{\text{èmeème}}$ siècle —de laquelle on peut dire, sans exagération, qu'elle est un développement des idées de Riemann— de la philosophie dans les mathématiques. Ce sont les mathématiciens, les Hermann Weyl , les Elie Cartan , qui formulent les conditions précises du problème et avancent, à partir des solutions proposées, les hypothèses interprétatives concernant les grandes questions qu'induit le développement de la relativité : structure causale de la relativité (l'espace-temps), interprétation des principes d'invariance, théorie de la matière. . . , tant il est vrai que le travail mathématique conduit ici, par une sorte de nécessité, à une théorisation physique.

On n'a pas cherché à traduire systématiquement les concepts en jeu dans un langage mathématique moins archaïque, notamment celui des espaces fibrés mis au point, dans les années 1950, par Charles Ehresmann, *a fortiori* dans d'autres qui pourraient être jugés plus actuels encore. Il y a à cela deux raisons principales, dont nous avons parfaitement conscience qu'elles ne parviendront pas nécessairement à faire recevoir comme excuse l'affirmation du fait que ce double archaïsme est pleinement assumé.

La première est que le langage de Cartan et Weyl [1] garde, dans sa proximité aux situations géométriques, une capacité d'évocation intuitive assez remarquable pour justifier de ne pas le changer.

La seconde est de fidélité historique. Quels que soient les efforts du " traducteur ", une partie de ce qui fait l'originalité et la force de ces mémoires, comme source mathématique encore vivante, est perdue dans la traduction.

1 De Kant à Riemann : l'horizon philosophique du problème.

1.1 Le problème de la géométrisation comme problème transcendantal.

1.1.1 La conception kantienne de l'espace

Pour Kant, la spatialité, c'est-à-dire l'espace euclidien, constitue, avec le nombre, c'est-à-dire le nombre entier, la forme de toute pensée d'objet. L'espace a une fonction trans-

[1] Nous nous limiterons en fait, par souci de brièveté, à l'œuvre d'Elie Cartan.

cendantale au sens où il s'agit, comme pour le nombre, d'un concept constitutif de toute représentation d'un monde d'objets : comment penser la droite, et le continu en général, sans le nombre réel, sans cette médiation essentielle qu'est la droite numérique ? L'espace est une condition *a priori* de la pensée des objets du monde extérieur. Les objets " du sens externe " sont toujours dans l'espace, ce sont toujours des objets spatiaux, des phénomènes. Et cette condition est une intuition, antérieure à toute représentation sensible d'un objet, donc *a priori*, plus exactement une des formes de l'intuition *a priori*, l'autre étant celle du sens interne, le temps. L'intuition de l'espace est en même temps la condition qui permet à la géométrie d'établir les propriétés de ses objets dans l'espace, par une construction synthétique, qui diffère du tout au tout du mode de connaissance, purement conceptuel, que Kant attribue à la philosophie.

L'espace n'est donc pas à seulement, ni même peut-être à proprement parler, un objet de pensée, mais une condition (*a priori*) de la pensée d'objet —de *toute* pensée d'objet. Chez Kant, la spatialité est, peut-on dire, à la fois *objet* de pensée et *condition* de pensée.

Elle est condition, d'abord, bien sûr, puisque c'est de là qu'il tire son caractère transcendantal.[2] Une réflexion transcendantale établit les conditions de la possibilité de considérer une entité comme objet, donc de connaître un objet —conditions *a priori*, mais, s'agissant de mathématiques, *l'a priori* va de soi. La spatialité est un schéma de la pensée objective, qui détermine *a priori* le cadre de toute pensée d'objet, source à la fois de l'espace perceptif et de l'espace mathématique, dans la mesure où en tant que système de règles, la spatialité peut se trouver projetée dans le monde perceptif et mathématique. . . [3]

L'espace est aussi objet de pensée : la mathématique constitue la spatialité comme objet opératoirement constitué. En ce sens, elle traduirait le schématisme kantien, mais un schématisme non autonome, lié à des systèmes d'opérations. C'est bien un objet, et un objet dont l'unité est *synthétique* —synthétique *a priori*. Un " concept naturel ", au sens que G. Granger a donné à cette expression [4], a une unité synthétique, en ce qu'il associe de façon nécessaire des propriétés non déductibles les unes des autres, qu'il réalise la synthèse entre des éléments ensuite dissociés par l'activité mathématique. Aux prises avec des contraintes et des obstacles nés précisément de cette solidarité des éléments constitutifs du concept, qui s'oppose à la mise en évidence de ces éléments, le travail mathématique doit dissocier ces concepts naturellement synthétiques en concepts partiels.

En tant qu'objet constitué de manière synthétique, le concept d'espace manifeste une intrication essentielle. Tous les traits qui seront ultérieurement dégagés par l'analyse mathématique, par exemple algébrique et différentiel —exemples de l'analyse tensorielle, de l'algèbre multi-

[2]Les choses sont d'ailleurs un peu plus compliquées d'un strict point de vue kantien, puisque, selon Kant, l'espace représenté comme objet, comme cela a lieu en géométrie, contient plus que la simple forme de l'intuition, mais encore l'intuition formelle qui résulte d'une synthèse, laquelle apporte l'unité au divers fourni *a priori* par la forme.

[3]Les objets mathématiques se déploient en droit indépendamment de tout contenu sensible —et même de ce que Kant appelait les formes de l'intuition sensible, car ce sont des systèmes de symboles et on les définirait adéquatement, propose G. Granger [3], plutôt que comme des formes d'intuition *a priori*, comme des formes *a priori* d'expression : des sortes d'"universaux du langage", fondant à la fois l'organisation d'un monde en objets sensibles et la production des objets mathématiques. Il faut reconnaître, contre Kant, qui voulait restreindre les schèmes formels de l'espace et du nombre aux intuitions sensibles, que leur application s'étend au delà de la constitution d'un monde d'objets sensibles. Il n'y a pas reproduction de certaines formes de l'intuition, mais système de contraintes déterminant les formes de l'intuition.

[4]Op. cit. note précédente ; voir aussi [4]

linéaire, ou algébrique et topologique —exemple de la théorie homologique de la dimension, sont inséparablement liés dans le concept originaire de spatialité. C'est ce qui fait que, même dans la pure mathématique, l'idée d'espace a une consistance, exige la concaténation de plusieurs points de vue.

En bref, l'espace serait pensable comme un *méta-objet*[5], un objet-condition, condition d'objets, d'autres objets. L'objet lui-même est condition de représentation de différents objets spécifiques : d'une part, objet plus général, plus abstrait, objet-forme ; d'autre part, condition de possibilité : en ce sens, non situé tout à fait sur le même plan, objet " de deuxième intention " auraient dit les scolastiques.

C'est de cette constitution de la spatialité comme méta-objet en général que vient la profondeur de la conception kantienne, devenue ensuite référence des grands mathématiciens (Gauss , Riemann, Poincaré, H. Weyl) et point de départ d'une conception renouvelée de l'espace comme concept mathématique. C'est ce qu'on pourrait appeler la " réflexivité " de l'espace. Dire qu'on peut concevoir l'espace comme une production de la pensée réflexive, c'est précisément reconnaître qu'il est, en même temps qu'objet de pensée, condition de la pensée d'objet. Ce qui était condition devient objet : il y a déplacement thématique, ce que nous pouvons appeler, après Kant, mais en un sens différent, " réflexion ".

On retrouverait là le problème particulier du statut de la géométrie, une science qui prend pour thème des objets de deuxième ordre, c'est-à-dire des systèmes théoriques (ou symbolico-théoriques) référés à des objets, plutôt que ces objets eux-mêmes : une discipline *réflexive*, qui n'a pas à proprement parler d'objets, ou dont les objets sont " de seconde intention ". Cette réflexivité peut aujourd'hui s'étendre aux mathématiques entières, dans la mesure où ces dernières ne cessent de produire, à l'intérieur d'elles-mêmes, des régions ou domaines d'énoncés, qui, s'ils portent originellement sur des objets, des objets de connaissance mathématique, sont susceptibles de devenir eux-mêmes, au terme d'un déplacement thématique, des objets : des êtres mathématiques, posés avec leurs règles opératoires —d'où sont nées, notamment, les structures, sans qu'on puisse pour autant en borner ainsi l'origine.

1. 1. 2. La position kantienne du problème de la géométrisation.

C'est également chez Kant qu'on trouverait, selon nous, les origines de ce qu'on pourrait appeler le " problème de la géométrisation ".

Le cœur de la doctrine kantienne de la connaissance, c'est la synthèse *a priori* de l'expérience. Mais, comme l'a montré Jules Vuillemin [5], ch. 1, §1., pour rester fidèle au projet critique qui comporte comme visée d'essence une reconnaissance des sciences constituées, Kant a dû admettre deux types de synthèses : la synthèse proprement mathématique, et la synthèse physique. La première concerne l'intuition ; la seconde, l'existence. La synthèse mathématique, qui ne donne que la possibilité de l'objet, est synthèse de l'homogène : c'est elle qui, par exemple, produit le continu linéaire. La synthèse physique l'est de l'hétérogène, mais elle dépasse la simple *possibilité* de l'objet pour le poser dans l'*existence*. Seule la physique donne la possibilité complète —intuition et existence— de l'objet de l'expérience, donc

[5]Le devenir de la physique moderne a montré que ce méta-objet peut avoir une fonction physique aussi bien qu'un objet de la tradition ; ainsi, la spatialité du mathématicien peut servir à une physique dépassant le donné immédiat. Par exemple, l'" espace de Hilbert " donne l'exemple d'une structuration mathématique très éloignée de notre échelle : on peut y avoir une infinité de dimensions. On obtient un concept très différent, à la fois de l'espace euclidien et de l'espace non-euclidien, et pourtant on continue d'appeler une telle structure, "espace", et cet espace a un usage physique, il sert au physicien à une autre échelle.

de l'expérience, thème de la réflexion transcendantale. Ainsi, il est bien différent d'effectuer une synthèse spatiale : par exemple de réunir deux points d'une ligne droite en construisant cette dernière dans l'intuition, comme le dit si bien Kant, et d'effectuer une synthèse proprement physique, comme celle qui engendre la relation de causalité, laquelle relie des entités ayant une existence physique, c'est-à-dire matérielle déterminée. On n'a plus alors d'évidence intuitive et directe de l'objet, mais seulement une certitude, discursive et indirecte. On pense l'objet sans le construire, on le rencontre sans le créer. D'où le fait que nos principes se trouvent alors affectés d'une irréductible contingence. Il faut en effet, à titre de condition originaire, la matière, le choc premier du donné de la sensation, et ce fait n'est pas sans soulever de sérieux problèmes à l'intérieur même de la doctrine[6].

C'est à ce niveau que se pose le véritable problème transcendantal : celui de l'articulation de la mathématique, de la géométrie, à la physique, de l'espace-objet, et objet d'une science indépendante, la géométrie, à l'espace-condition, condition de constitution de l'intuition empirique, et forme du monde —le monde matériellement existant de la physique.

Dans le cadre de la Critique, il s'agit de produire la suite des médiations et des synthèses qui acheminent de la synthèse mathématique à la synthèse physique. C'est tout l'enjeu de ce que Kant appelle le schématisme des concepts de l'entendement : la constitution progressive, par une sorte d'approximation continuée, de l'expérience, avec la position successive des différents donnés.

Il n'est pas nécessaire d'entrer dans les détails de cette construction, au demeurant affectée de nombreuses incertitudes, notamment dans la jonction qu'il lui faut bien opérer, à un moment donné, avec la connaissance physique proprement dite. L'important pour nous est ici de remarquer que le véritable problème transcendantal naît au point exact de l'articulation de la mathématique à la physique —de la synthèse mathématique à la synthèse physique. Lié à la physique de son temps [7], Kant conçoit cette articulation sur le mode de la mise en équation, au sens que ce terme revêt dans la physique post-cartésienne et même post-newtonienne, celle qu'illustrent par exemple d'Alembert, Euler, Clairaut, etc. La situation physique est susceptible d'une expression par des équations différentielles. Les exemples en seraient nombreux durant tout le XVII$^{\text{ème}}$siècle, que l'on pense seulement à l'équation des cordes vibrantes. Aujourd'hui, la situation est différente. Quand on cherche à poser le " problème de l'espace ", il s'agit plutôt de géométrisation. Quel sens précis donner à ces deux expressions ? On ne peut répondre sans référence à l'histoire, car la position même de ces problèmes est la conséquence d'une double événement, le premier, dans l'histoire des mathématiques, le second, dans l'histoire de la physique : la géométrie de Riemann d'abord, la théorie einsteinienne de la relativité, ensuite.

[6] C'est la problématique générale des post-kantiens. N'y aurait-il pas, entre le réel et notre entendement, un *abîme* qui ne peut être comblé ? Et Kant n'est-il pas obligé d'admettre finalement un " hiatus " , selon sa propre expression dans la *Critique du jugement*, entre l'analogie universelle, constituée par les lois pures de l'entendement transcendantal, et les analogies particulières, ou empiriques, que sont les lois de nos sciences ? C'est le problème de l'*applicabilité* des formes à la matière : y a-t-il dans les choses hors de nous un fondement, une disposition quelconque, qui leur rende applicables les formes sensibles subjectives de notre esprit ? Kant aborde le problème dans la *Critique du Jugement*.

[7] C'est, rappelons-le, celle d'Euler plus précisément que celle de Newton, comme on le dit trop volontiers, même s'il est vrai que la première s'inscrit naturellement dans l'héritage de la seconde. Il y aurait évidemment beaucoup à dire sur cette dernière, envisagée du point de vue adopté ici, de la géométrisation.

1.2 La position riemannienne du " problème de la géométrisation "

On a remarqué que le point de vue de Riemann, dans sa thèse d'habilitation de 1854, procédait au fond un point de vue kantien. Il consiste en effet à se demander quelles sont les conditions de possibilité d'une mesure de grandeur dans l'espace, opération apparemment considérée par Riemann comme fondatrice de toute géométrie.

Une première ambiguïté concerne ce qu'on appelle exactement " espace ". Aujourd'hui, compte tenu des développements mathématiques postérieurs, que l'œuvre de Riemann a, pour une grande part, suscités, on entend par là un être abstrait, défini par des conditions très générales, notamment la possibilité d'y déterminer un repérage par des fonctions coordonnées. C'est le concept de variété, construit par l'intermédiaire de cartes et d'atlas. Riemann entend par là au contraire, semble-t-il, l'espace, *der Raum*, dit-il, et il semble bien désigner ainsi une entité unique (" singulière " disait Kant d'où son caractère intuitif) —le milieu de l'expérience, le site des corps et le lieu du mouvement. C'est l'espace de la tradition physique, en tant qu'il est investi par la géométrie euclidienne, un réservoir de points dont l'existence donnée garantit l'applicabilité des relations physiques : voir par exemple en mécanique la notion de masse ponctuelle [8].

L'événement dont il convient de tenir compte, insoupçonné de Kant sous sa forme radicale, c'est la pluralité des géométries. Le " problème de l'espace " devient la question de savoir laquelle de toutes ces géométries est vraie de *l'*espace.

L'originalité de Riemann est ici, on le sait, de prendre la question au niveau le plus général : l'espace n'y est qu'une détermination à l'intérieur du genre plus général de la grandeur étendue, dont chaque espèce est caractérisée par une géométrie [9]. Cela revient à introduire (sans que Riemann dispose d'une terminologie appropriée, *a fortiori* de celle qui deviendra usuelle) ce qu'on appelle aujourd'hui une espace topologique : l'espace de l'*analysis situ*s, dit Riemann, avec la propriété supplémentaire que chacun de ses points a un voisinage homéomorphe à une région de R^n pour une valeur donnée de n[10].

La grandeur qui forme le support de la mesure est elle-même considérée sous l'aspect le plus général, puisqu'il s'agit d'une grandeur quelconque à n dimensions, qu'il appelle " multiplicité ", par exemple la grandeur étendue est une multiplicité à 3 dimensions. Supposant que la valeur d'une telle grandeur est indépendante de sa position (*Lage*) sur la multiplicité qui la contient, Riemann se propose de formuler les conditions de sa mesure, cette dernière

[8] C'est en ce sens que le prend aussi Gauss , envisageant de vérifier par des mesures géodésiques la validité empirique des principes de la géométrie euclidienne.

[9] Riemann hérite ici de Gauss , qui déclare dans ses Contributions à la théorie des équations algébriques, [6] : " A strictement parler, le contenu essentiel de l'argument appartient à un domaine plus élevé, indépendant de l'espace, de la science abstraite de la grandeur, où l'on recherche les combinaisons de grandeurs reliées entre elles (zusammenhängenden Grössencombinationen) par continuité. Ce domaine est encore peu développé, et on ne peut en traiter avec le langage emprunté aux images spatiales. " Le point de départ est le même que chez Riemann : une structure qu'on dirait aujourd'hui de variété, munie de systèmes de coordonnées par des continus de n-uples de nombres, qui permet d'user d'un langage géométrique en contexte non géométrique. Là où Gauss introduit ensuite une structure algébrique, celle des complexes, Riemann introduit des n-uples de réels et une structure métrique.

[10] Il s'agit en fait de " variétés ". L'objet dont part Riemann, tant dans ses recherches sur les fonctions de variable complexe que dans ses recherches de géométrie, c'est la variété, plus précisément, la variété munie d'une structure analytique : structure complexe dans le cas des surfaces de Riemann, différentiable dans le cas de la géométrie différentielle. Riemann ne l'interprète pas seulement au sens de la topologie ensembliste, mais encore en celui de la topologie combinatoire, considèrant des relations de " connectivité ".

étant définie en général comme report d'une ligne sur une autre, par exemple la mesure des longueurs des lignes (“ multiplicités de dimension 1 ”) tracées sur une surface (“ multiplicité de dimension 2 ”), et en général l'évaluation des multiplicités à n dimensions dans les multiplicités à n+1 dimensions. Le problème de la recherche des conditions de possibilité se trouve déterminé, plus précisément que chez Kant, comme celui des axiomes qui sont susceptibles d'expliciter ces conditions dans une géométrie donnée. Ces axiomes ne peuvent être qu'*a priori* au sens de Kant : pour les formuler, il est nécessaire de dépasser le champ de l'expérience, en faisant des hypothèses relatives notamment à l'infiniment petit. La question préalable posée par Riemann est celle de savoir si ces hypothèses découlent nécessairement de la géométrie donnée, ou si elles sont simplement compatibles avec elle. On connaît sa réponse, imposée par la pluralité des géométries. Il n'y a pas ici de détermination nécessaire. Le genre “ grandeur trois fois étendue ” peut être spécifié de plusieurs manières, car il y a plusieurs modes possibles de détermination de la distance, donc des rapports entre les paires de grandeurs spatiales homogènes, d'où provient la géométrie. Il y a plusieurs relations métriques, ou métriques, alternatives, et ainsi plusieurs géométries compatibles, entre lesquelles, aux yeux de Riemann, seule l'expérience permet de décider.

Comme celui de l'espace, le “ problème de la géométrisation ” n'a pas de sens vraiment précis avant Riemann, en entendant par “ sens précis ” une mathématisation réussie. Dans le mémoire de Riemann, il s'agit de s'interroger sur “ les faits les plus simples ” qui permettent de déterminer les relations métriques de l'espace. Une fois admise et reconnue la pluralité des géométries, on ne peut plus retenir l'idée d'une déduction nécessaire de tels faits à partir de concepts généraux de la mesure. Point ici d'énoncés nécessaires, mais seulement des hypothèses, dont la signification est de simple probabilité. Ultime leçon de Kant, l'intervention de l'expérience est inéluctable, la légitimation ne peut être ici purement conceptuelle, mais en appelle à l'empirie.

1.3 Nature philosophique des prises de position quant au “ problème de l'espace ”

L'œuvre mathématique qui montre le plus clairement à quel point l'analyse mathématique de l'espace est inséparable de prises de position d'ordre philosophique est sans doute celle de Poincaré. Quelque quarante ans après la parution du mémoire de Riemann, introduisant en France, à la suite de ses propres travaux mathématiques, la conception de Klein de réduction de la géométrie à l'étude d'un groupe, Poincaré se situe, lui aussi, dans l'héritage kantien. Le refus de la solution “ rationaliste ” d'une forme *a priori* de notre sensibilité vient directement, plus nettement encore que chez Riemann, de la reconnaissance de l'égalité de statut entre les géométries euclidiennes et non-euclidiennes. Mais Poincaré ne suit pas Riemann jusqu'au bout. Le “ conventionnalisme géométrique ” est présenté comme un intermédiaire entre l' “ empirisme géométrique ” , qu'il décèle chez Riemann, et qu'il critique, et le rationalisme kantien, qui lui paraît à réformer compte tenu des développements de la géométrie.

On connaît le principe de la solution du “ conventionnalisme géométrique ”. L'établissement d'une relation, ou d'une connexion, précise entre expérience et géométrie spatiale est affaire de *conventions*. On impose une géométrie à notre espace par le moyen de postulats dont le statut n'est que celui d'énoncés conventionnels, quoi que ce soit l'*expérience* qui nous guide jusqu'à eux. On choisit ceux qui sont les plus appropriés à l'expérience et on leur confère

le statut de principes agréés, les protégeant ainsi d'une réfutation future. Les conventions ne sont pas empiriquement contrôlables, bien qu'on puisse concevoir que, pour de bonnes raisons, notamment liées à l'expérience, on puisse remplacer une convention agréée par une autre.

Or, ce point de vue de Poincaré procède d'hypothèses épistémologiques fondamentales. L'une de celles-ci porte précisément sur le statut épistémologique de l'espace. Selon lui, nous ne connaissons pas l'espace directement. On ne peut percevoir, ni appréhender de quelque façon que ce soit, quelque chose comme l'espace, par où Poincaré entend alors évidemment par là, comme Riemann, cette entité absolue postulée par la mécanique et la physique de son temps. C'est toujours, au fond, sous divers revêtements, l'espace absolu de Newton. A partir d'une position subjectiviste (il adopte très tôt une psychologie associationniste et " sensationniste ", qui domine sa théorie de la genèse de la géométrie), Poincaré décrit la manière dont on acquiert les relations spatiales, et dont on crée l'espace géométrique. Ce qu'il est en notre pouvoir de créer, c'est un ensemble de géométries, ou de langages géométriques alternatifs. On choisit alors par convention l'un d'entre eux comme la meilleure représentation de notre espace, et le meilleur langage mathématique pour nos théories physiques[11].

Ainsi, alors que Riemann semble bien se référer à *l'*espace au sens de l'espace physique, source des points matériels qui forme le support de la réduction physique, Poincaré l'ignore complètement. L'expression même d' " espace physique " ne fait pas partie de son vocabulaire philosophique. De son point de vue, il y a un espace *sensible ou représentatif* construit à partir de sensations visuelles, tactiles, de mouvement, par un processus de classification[12]. Cependant, cet espace empirique n'a aucune des propriétés du véritable espace géométrique, n'étant ni infini (ou même non-borné), ni homogène, ni isotrope. Le propos de Poincaré reste fondamentalement épistémologique : il est moins d'expliquer l'origine et les propriétés de l'espace empirique, dans une psychologie, que celles de l'espace géométrique pur, espace conceptuel sur lequel nous raisonnons et que nous pouvons utiliser dans nos théories physiques. C'est le concept relativiste, en un sens qui n'est pas exactement celui d'Einstein, de l'espace.

2 L'algébrisation du problème : l'influence de F. Klein et du " Programme d'Erlangen "

La géométrie riemannienne est le premier exemple d'un problème de géométrisation au sens précis qu'on donnera ensuite à cette expression. Il s'agit de géométrie différentielle, puisqu'elle est définie par la donnée de la distance entre deux points infiniment proches. La distance est donnée par la racine carrée d'une forme différentielle quadratique (définie positive) dont les coefficients sont des fonctions (supposées en général indéfiniment dérivables)

[11] Poincaré plaide notamment pour la possibilité de deux systèmes alternatifs, formulables en thèses de " philosophie du non " à la manière de Bachelard : 1) non-euclidianisme : la géométrie spatiale peut être une forme de géométrie non-euclidienne ; 2) non-tridimensionalité : il est possible de concevoir notre espace autrement qu'à trois dimensions.

[12] Poincaré considère cependant les sensations comme ayant une fonction significative dans la création et le développement de notre concept d'espace géométrique, dont elles fournissent à notre esprit l'occasion de la construction. Notre esprit forge une relation entre les sensations et l'espace géométrique par des actes de classification et d'analyse de certains changements de nos sensations.

des coordonnées x^i, appelée encore " forme fondamentale " ou " ds^2 " de l'espace. On obtient ce qu'on appelle une " variété ". La géométrie euclidienne rentre dans ce schéma très général : avec des coordonnées cartésiennes, la forme fondamentale y est à coefficients constants.

On sait qu'on peut développer cette géométrie en partant *seulement* de cette forme. En effet, on établit que le groupe des déplacements et retournements est *complètement* caractérisé par la propriété de laisser invariante la forme fondamentale. Donc on obtient, en même temps que les propriétés géométriques, toutes celles qui s'en déduisent par changement de coordonnées. Il en va de même en géométrie riemannienne, dont on peut dire qu'elle est l'étude des propriétés communes à une variété V et à toutes les variétés qui lui sont " équivalentes dans la géométrie du groupe G ".

Dans tous les cas, on voit qu'il s'agit de rattacher, à une forme différentielle quadratique, des propriétés géométriques indépendantes du choix des variables c'est-à-dire invariantes par changement de coordonnées. Avec l'invariance, c'est l'intervention inéluctable du concept de groupe, qui nous renvoie à F. Klein et au " Programme d'Erlangen ".

Comme le dira Weyl , ce dernier marque le triomphe de la plate-forme démocratique en géométrie. Avec la pluralité des géométries possibles, le problème se posait d'une *description générale* de la *totalité* des systèmes géométriques. L'idée fondamentale de Klein est de considérer les différentes géométries comme théories d'invariants de groupes appropriés. Une fois fixée une certaine famille C de transformations, une géométrie se définit comme l'étude des propriétés des objets —des figures— qui sont conservées par ces transformations. Une géométrie est ainsi déterminée par deux éléments : un domaine d'action E, ou " espace ", et un groupe de transformations du domaine ou de l'espace dans lui-même, plus précisément d'automorphismes, formant un groupe, le " groupe de symétrie ", qui opère sur ce domaine E, en laissant certaines propriétés ou relations invariantes. Ainsi, la géométrie dite *euclidienne* peut-elle se réduire d'une part à des objets, figures planes et spatiales (ensembles de points du domaine, plan ou espace), d'autre part à un groupe de transformations, ou isométries, qui transforment les figures en d'autres " égales " aux premières, puisqu'elles laissent invariantes les distances entre leurs points respectifs. La géométrie euclidienne se trouve définie comme l'étude des propriétés invariantes par isométries (transformations conservant la distance entre les points). D'où un nouvel usage de l'expression " géométrique ", réservée dorénavant à toute propriété des figures du domaine et à toute grandeur liée à ces figures telles qu'elles soient invariantes par toute transformation appartenant au groupe donné. C'est le système des propositions relatives aux propriétés des figures et grandeurs invariantes par toutes les transformations du groupe qui constitue la géométrie de ce groupe. Relativement à Euclide , le terme " géométrie " a changé profondément de sens. Une " grandeur géométrique " n'est plus une quantité, déterminée plus ou moins intuitivement dans l'espace par la considération directe des figures, ou par leur mesure, c'est l'invariant d'un groupe de transformations. Changeant le groupe, on change le schéma géométrique, on obtient une " nouvelle géométrie ". La dualité fondamentale d'un espace, comme contenant dénué de toute propriété autre que topologique, et d'un groupe, comme fondement de la notion d'égalité, source de toutes les propriétés géométriques des figures, c'est selon Poincaré celle de la " matière " et de la " forme " : la matière, c'est le champ opératoire, l'ensemble des objets, en nombre fini ou infini sur lesquels opèrent les transformations du groupe ; la forme, c'est la loi de composition ou " structure " du groupe, la loi de formation de l'inverse

d'un transformation et du produit de deux transformations : ce qu'il y a d'essentiel dans le groupe quand on le dépouille de ce qu'il a de contingent, c'est-à-dire tenant à la nature des objets sur lesquels il opère.

Dans une telle perspective, il est difficile de voir dans l'espace autre chose qu'un contenant, dépourvu de tout contenu, de toute spécification autre que topologique. Tout se passe comme si les anciens attribuaient à l'espace des propriétés qui étaient en fait celles des figures de l'espace, ou plongées dans l'espace. Avec Klein, le point de vue change : on considère que les propriétés projectives, par exemple, ne sont pas celles des figures mais celles de l'espace lui-même, en tant qu'il est spécifié par la donnée du groupe : l'espace projectif, comme champ d'opération des transformations du groupe projectif. C'est l'espace qui supporte le génitif. Il est sujet, et non plus objet, du génitif. On parlera de l'espace *de* tel ou tel groupe, et non pas du groupe de transformations *de* l'espace. C'est l'apparition de ce qu'on a nommé la réflexivité : remontée de la pensée de l'objet à la condition de celle-ci, encore mathématiquement déterminée.

Cette montée vers l'abstraction donne prise, dans un premier temps aux interprétations " conventionnalistes ". Les déterminations très générales ne peuvent être introduites que par convention, règles ou axiomes. On attribue, par commodité de langage, à l'espace dans lequel sont localisées les figures, les propriétés du groupe lui-même. Il y a projection du groupe dans son champ d'action, l'espace sur lequel il opère. D'où résulte cette abstraction : la géométrie du groupe. Ainsi, quand j'énonce la propriété, fondamentale en géométrie, d'indépendance des figures relativement à leur position dans l'espace sous la forme de l'axiome " deux objets égaux à un même troisième sont égaux entre eux ", je projette dans l'espace la propriété des déplacements de former un groupe. L'axiome représente en quelque sorte la version ontologique de la formulation épistémologique, qui relève de la théorie des groupes[13].

Le nouveau point de vue imposé par Klein montre que la vieille géométrie euclidienne, prise dans sa totalité, synthétisait en fait deux caractères, ultérieurement dissociés : son caractère fondamentalement métrique, et l'homogénéité.

Le premier est évident, et justifie pleinement qu'on ait appelé " euclidien " un espace dans lequel on peut définir une métrique —même si les propositions démontrées par Euclide ne sont pas exclusivement de nature métrique. Le second signifie que les propriétés sont les mêmes en tout point du domaine, ou de " l'espace ", dans lequel opère le groupe de transformations considéré. Ce n'est qu'une version de l'invariance : cela revient à dire que les propriétés sont invariantes par les transformations du groupe correspondant.

On sait qu'on a fini par appeler " espaces homogènes " les espaces de groupes tels que tout élément est le transformé d'un autre par un élément du groupe —un couple (E,G) où G est un groupe (dit " d'opérateurs ", avec une loi de composition interne) opérant transitivement dans E[14]. Dire que tous les espaces considérés par Klein sont homogènes, c'est poser que

[13]On constaterait la même évolution en ce qui concerne la structure. E. Cartan commence à parler de " structure ", mais, lorsqu'il le fait, c'est toujours avec le génitif : on parle de structure *de* tel ou tel groupe, comme dans le passage suivant : " il subsiste cependant quelque chose de commun entre les deux formes de la géométrie projective (distinguées par la donnée de l'élément générateur, droite et non pas point), et c'est la structure du groupe, c'est à dire la loi de composition de ses opérations... " [7].

[14]Le groupe G " opère " dans E s'il existe une loi de composition externe, application de GxE dans E telle que si x∈E, alors, pour tout couple (g,h) d'éléments de G, on ait : 1) g(hx)=(g*h)x ; 2) e étant élément neutre de G, ex=x. Il opère " tnasitivement " si, pour tout couple d'éléments (x,y) de E, il existe un élément g de G tel que y=gx.

toutes les propriétés en sont invariantes par les transformations de ce que Klein appelle le " groupe fondamental —une dénomination tombée ensuite en désuétude, au profit de la notion homonyme de topologie algébrique ; on pourrait dire " groupe caractéristique ". Le sens en est plus concrètement géométrique.

Prenons l'exemple de l'espace des transformations projectives. On peut démontrer que tout invariant de l'espace projectif peut s'exprimer en termes d'un invariant unique, le rapport anharmonique. Donc tout groupe de transformations qui est un groupe d'invariance, c'est-à-dire dont les transformations laissent une certaine relation invariante dans l'espace projectif, est un sous-groupe du groupe des transformations laissant invariant le rapport anharmonique, c'est-à-dire du groupe projectif. Ce dernier peut être à bon droit appelé " groupe fondamental " de l'espace projectif, au sens où tout autre groupe d'invariance est un de ses sous-groupes. Le groupe fondamental chez Klein sert en quelque sorte de mesure de l'homogénéité des espaces dans lesquels ces groupes opèrent. Il y a une sorte d'équivalence entre tous les objets d'un certain point de vue : celui qu'exprime la transformation considérée.

On a donc deux situations extrêmes, entre lesquels on trouve tous les degrés mathématiquement concevables. A un pôle, un espace parfaitement homogène serait un espace dans lequel le groupe fondamental serait le groupe de toutes les transformations continues. Il coïnciderait avec l'espace métriquement amorphe de la topologie. Ses propriétés géométriques en sont peu variées et peu intéressantes. A l'autre pôle, un espace complètement dépourvu d'homogénéité serait un espace dont le groupe fondamental consisterait dans la seule transformation identique. Dans un tel espace, évoquant le monde des anciens mégariques, on ne pourrait sortir de la considération de cas, d'entités singulières ou d'instances individuelles d'objets. Comme le dit Cartan , il n'y aurait alors pas de science du général. Toute la géométrie deviendrait un chaos de faits particuliers, sans possibilité de mise en relation. Il n'y aurait même, s'il est vrai, comme l'affirme Aristote, qu'il n'y a de science que du général, pas de connaissance objective possible.

Il convient de remarquer que cette propriété d'homogénéité met le point de vue de Klein dans une divergence radicale d'avec celui de Riemann.

Fondée sur un concept métrique, même très large (mesure d'une multiplicité, longueur d'une ligne) cette dernière ne rentre pas dans le programme d'Erlangen, dont le point de vue apparaissait pourtant de la plus grande généralité. Dans l'espace riemannien, on ne soumet pas le concept de mesure à des contraintes *a priori*, sous la forme de lois universelles. On ne fait pas intervenir, dans toute région de l'espace, l'espace entier. Une mesure comme la longueur est supposée définie de proche en proche par une forme différentielle, généralement quadratique, mais en fait arbitraire. Il n'y a plus renvoi au global, on se meut partout dans du pur local. Dans un tel espace, on n'a pas d'homogénéité, ou plutôt, il n'y a d'homogénéité qu'infinitésimale. On peut identifier l'espace riemannien à un espace euclidien au voisinage d'un point, mais on ne peut aller au-delà. Il n'y a plus de liaison entre les différents espaces euclidiens ainsi distingués.

Une conséquence est qu'on ne peut appliquer un principe mis en évidence par Klein dans le cadre du " programme d'Erlangen, fondement de sa classification des géométries, et qui semblait d'application universelle : le " principe d'adjonction ", qu'on exprime souvent de la manière suivante :

Si V est une variété n-dimensionnelle et H un groupe de transformations de V, on définit une géométrie sur V en " adjoignant " H à V. Une telle géométrie consiste dans la théorie

des H-invariants. Si H est un sous-groupe d'un groupe donné G, on définit une nouvelle géométrie sur H à partir de la précédente sur G.

On peut présenter les choses, de manière équivalente, mais peut-être plus concrète, d'un autre point de vue, à partir des figures, plus généralement des relations, laissées invariantes par le groupe ou sous-groupe en question. Si les transformations adjointes sont celles qui laissent une certaine figure, dite " distinguée ", ou " invariante ", l'opération revient à adjoindre cette figure aux anciennes. Ainsi peut-on dériver la géométrie affine de la géométrie projective, en *adjoignant* à l'espace projectif un plan distingué, qui est le plan à l'infini. Le groupe fondamental de la géométrie affine est un sous-groupe du groupe fondamental de la géométrie projective, et les transformations de ce sous-groupe sont celles qui laissent invariant le plan à l'infini. De même, l'espace de la géométrie euclidienne peut être obtenu à partir de l'espace affine ou projectif en y adjoignant une conique privilégiée, ou distinguée, située dans le plan de l'infini, et qui est le cercle imaginaire à l'infini. Quant aux espaces non-euclidiens (" lobatchevskiens " ou " riemanniens ") ils sont obtenus par distinction d'une quadrique non dégénérée, qu'on appelle " l'absolu ", etc. D'où un principe général, dit " de subordination ", qui est aussi un principe de classement : si H est un sous-groupe de G, alors la théorie des H-invariants est une autre géométrie, subordonnée à la première. Ainsi les diverses géométries : élémentaire, affine, non-euclidienne, sont-elles subordonnées à la géométrie projective.

On pourrait alors, en poussant un tel point de vue à l'extrême, considérer que la géométrie riemannienne n'est rien d'autre que la géométrie du groupe infini des transformations ponctuelles à n variables, auquel on a adjoint une forme différentielle quadratique déterminée. Elie Cartan remarque que cela reviendrait à priver de son sens le principe de Klein, en le conduisant à l'absurde. A ce compte en effet, on pourrait dire que tout problème géométrique rentre dans le " programme d'Erlangen " : il suffit d'adjoindre au groupe de toutes les transformations possibles les données du problème à résoudre [8] !

C'est par ce biais qu'on peut comprendre la relation nouvelle qui se noue alors entre la géométrie et la physique. Dans le champ de la physique, la théorie de la relativité introduisait en effet exactement la même disjonction que celle qui se manifestait en géométrie. Elle aboutissait d'ailleurs à séparer les deux théories successivement édifiées par Einstein. Les espaces-temps de la mécanique classique et de la Relativité restreinte sont homogènes ; ceux de la Relativité générale ne le sont pas : ils sont du type de ceux de Riemann, et, dans cette différence, se tient un enjeu essentiel pour la théorie physique, et sa géométrisation.

Avant d'y revenir, il nous faut évoquer la fonction de la relativité dans le champ de l'histoire de la géométrie et de la philosophie de l'espace.

3 Le rôle de la relativité : la synthèse géométrique d'E. Cartan

3.1 La leçon de la relativité

Hermann Weyl a souligné ce fait que l'influence de la théorie de la relativité a été dans tous les cas profonde, mais de nature bien différente selon qu'on l'entend sous l'une ou l'autre de ses deux formes. La restreinte a eu l'influence la plus forte sur la physique ; ce fut

le contraire pour la relativité générale, qui a eu l'influence la plus forte sur les mathématiques —de plus faible portée sur la théorie physique.

En Relativité restreinte, la notion de relativité exprime l'homogénéité de l'espace-temps considéré comme domaine d'action du groupe, dit " de Lorentz ", des transformations qui conservent le carré de l'intervalle élémentaire $ds^2=dt^2-\frac{1}{2}(dx^2+dy^2+dz^2)$. La géométrie à quatre dimensions ainsi définie est une géométrie euclidienne à forme fondamentale indéfinie.

La transposition du géométrique au physique est immédiate.

Du point de vue physique, l'homogénéité de l'espace-temps agit comme principe de causalité : des causes égales opérant en des régions différentes de l'espace-temps et à des moments différents du temps produisent des effets égaux. D'où la nouvelle détermination de la loi physique comme relation dont l'expression ne change pas quand on passe d'un observateur galiléen à un autre, c'est-à-dire quand on lui applique la transformation de Lorentz. C'est l'invariance, ou covariance, relativiste. On peut montrer que les transformations de Lorentz sont en fait les rotations de l'univers à 4 dimensions, rotations d'angle imaginaire si on définit un temps imaginaire, pour avoir un espace, dit " de Minkowski ", vraiment euclidien. D'où il résulte que les lois relativistes sont invariantes par le groupe des rotations dans l'espace-temps.

Il faut donc exprimer les lois de la physique de façon qu'elles ne changent pas lors de rotations de l'espace-temps, qui comprennent, outre les rotations de l'espace ordinaire, qui en constituent un sous-groupe, les translations dans ce même espace ordinaire, et les transformations relativistes du temps). Des entités physiques bien déterminées, comme les vecteurs, et de façon plus générale, les tenseurs, seront égaux pour tout observateur s'ils le sont pour un observateur galiléen : par exemple, l'impulsion et le vecteur d'onde d'univers.

En Relativité générale, on n'a plus de principe de causalité de ce type : l'espace-temps n'est plus homogène en ce sens. On ne peut donc plus soumettre, comme en physique classique, le concept de longueur à des lois données *a priori* qui font intervenir, dans toute région de l'espace, l'espace entier. La longueur est définie de proche en proche par une forme différentielle. L'espace-temps de la Relativité générale est un espace riemannien à ds^2 indéfini et à 4 dimensions, qui appartient à une classe particulière d'espaces riemanniens [15]. Cette classe est caractérisée par l'invariance par changement de coordonnées. C'est ce dernier fait qui permet à l'espace riemannien d'exprimer géométriquement la relativité, au sens physique d'indépendance du choix des variables dont on se sert pour représenter analytiquement les phénomènes. Et c'est la définition analytique de cette classe d'espaces riemanniens qui donne les équations de la gravitation dans le vide.

Tout espace riemannien étant caractérisé par son tenseur fondamental, g_{ij}, son tenseur de courbure R^k_{ijk} et ses dérivés, les équations de la gravitation ne doivent faire intervenir que

[15] Les exigences de la relativité restreinte avaient contraint Einstein à apporter deux modifications à la géométrie riemannienne. D'abord le carré d'élément d'arc n'est plus une forme quadratique positive, mais indéfinie : c'est la métrique de Minkowski, $ds^2 = -dx_0^2 + \Sigma dx_i^2$, i=1...3, avec x_O=ct. Ensuite, la dimension de la variété n'est plus 3 mais 4 (il faut ajouter le temps aux trois dimensions de l'espace). L'espace de la Relativité générale est une variété à quatre dimensions, localement minkowskienne, et non localement euclidienne comme les espaces riemanniens ordinaires (tout espace tangent est doté d'une métrique de Minkowski), alors que les espaces tangents des espaces riemanniens sont, dans des systèmes convenables de coordonnées, euclidiens, c'est-à-dire vérifient $ds^2 = \Sigma dx_i^2$. Le point important est que, dans les deux cas, la forme quadratique est non dégénérée, condition fondamentale qui rend possible l'application de l'analyse tensorielle.

ces dernières expressions. Dans un système de repères orthogonaux, on obtient un système de relations linéaires entre les 20 composantes du tenseur de courbure. La recherche d'une solution invariante par changement de coordonnées est un problème d'algèbre qui admet plusieurs solutions : celle adoptée par E fait intervenir le tenseur contracté $\mathrm{R}_{ij} = R^k_{ijk}$.

Il s'agit bien d'un problème de géométrisation, puisqu'il s'agit de rendre le traitement indépendant du choix des coordonnées. Le problème traité par Einstein est tout à fait équivalent à celui qui consiste, en géométrie euclidienne, à chercher la classe des surfaces contenant, en même temps qu'une surface, toutes celles qui en dérivent par déplacement. En Relativité Générale, le groupe des déplacements est remplacé par le groupe (infini) G des transformations ponctuelles exprimées en terme des 4 coordonnées x_i, i=1 , ..., 4, avec les transformations correspondantes des coefficients g_{ij} d'une forme différentielle quadratique (l'espace euclidien est remplacé par l'espace à 14 dimensions des coordonnées x^i et g_{ij}). Etant donnée une variété, on cherche toutes celles qui s'en déduisent par une transformation du groupe. La géométrie riemannienne est l'étude des propriétés communes à une variété et à toutes celles qui lui sont équivalentes dans la géométrie du groupe infini G.

3.2 Structure de la synthèse d'E. Cartan

En général, on peut considérer un système d'équations différentielles, ou d'équations aux dérivées partielles ou de formes différentielles. C'est ce qu'Oswald Veblen appelle un " objet géométrique ". On peut considérer la classe de cet objet et de tous ceux qui s'en déduisent par une transformation quelconque, une transformation d'un certain groupe, appliquée aux variables. L'étude des propriétés communes à tous les objets de cette classe est typiquement un problème " de géométrisation ". C'est pour aborder les problèmes les plus généraux de ce type qu'Elie Cartan a généralisé la méthode dite " du repère mobile ", inventée par Levi-Civita au terme d'une première généralisation du concept en usage en géométrie différentielle des surfaces dans l'espace euclidien [9].

L'outil mis au point par Levi-Civita, qui mobilise un mode de déplacement nommé " transport parallèle ", permet de pallier l'inhomogénéité de l'espace riemannien. Se laissant guider par la notion de parallélisme, mise au centre de son travail, Cartan peut exploiter la notion de développement le long d'une courbe paramétrée, adoptée comme fil conducteur. C'est ainsi qu'il a été conduit à sa définition de ce qu'on appelle aujourd'hui une " connexion ".

Il s'agit de compenser le double fait que les propriétés sont, d'une part purement locales, donc variables en général d'un point à un autre, d'autre part différentes selon les espaces, donc permettant de les caractériser. Il est le moyen d'introduire dans un espace riemannien tous les déplacements euclidiens relatifs à une région infiniment petite.

On part de la notion d'espace tangent (euclidien) en un point de la variété, par exemple l'espace riemannien d'origine, et l'on se donne une loi permettant de raccorder les espaces tangents en deux points infiniment voisins, ce qui permet de procéder ensuite de proche en proche, et de relier entre eux tous ces espaces, euclidiens et définis en tous les points de l'espace riemannien d'origine.

Soient deux repères en deux points de l'espace infiniment voisins. L'introduction du " transport parallèle " permet d'orienter les axes du deuxième repère relativement à ceux du premier, donc de fixer l'orientation. On sait en effet reconnaître le cas où deux vecteurs d'ori-

gine A et A' doivent être considérés comme parallèles, puisque la rotation élémentaire connue par la donnée du ds^2 fournit l'orientation mutuelle de deux repères infiniment proches, et permet de décider dans quel cas deux vecteurs d'origines infiniment proches sont parallèles. L'opération de " transport parallèle " ne permet donc pas seulement de relier deux vecteurs en deux points de l'espace, elle permet de *raccorder dans un seul et même espace* euclidien (l'espace tangent en ce point à la variété), deux " petits " morceaux de l'espace riemannien. Certes, ce raccord va dépendre, en général, du chemin suivi pour aller d'un point à l'autre, c'est le défaut d'" holonomie ". Il suffit de relier, aux infiniment petits du second ordre près, les repères ayant pour origines deux points infiniment voisins, et tous les vecteurs qui en sont issus. Ainsi, un arc de courbe C, avec tous les repères attachés à ses points, et les vecteurs qui en sont issus, pourront être incorporés, de proche en proche, dans l'espace tangent. La courbe C et toute la région immédiatement voisine viennent par " développement ", en analogie avec le cas d'une surface, se localiser en lui, et peuvent donc être considérées comme partie d'un seul et même espace euclidien. Ce dernier est sans doute fictif : c'est l'espace euclidien considéré abstraitement, un espace dont la forme fondamentale est la forme donnée, dans laquelle les coefficients g_{ij} sont constants, et qu'on rapporte à un repère d'origine A. Mais on peut transporter les principales notions de géométrie euclidienne, par exemple toutes celles de la théorie des courbes, à la géométrie riemannienne, avec la courbure et la torsion. Les géodésiques, lignes de minimum de distance entre deux point, seront les lignes de courbure nulle, généralisation des droites. On finira par étendre la plupart des théorèmes de géométrie euclidienne.

La notion analytique qui fonde le parallélisme est celle de différentielle absolue d'un vecteur mobile : ses composantes sont exprimées au moyen des symboles dits " de Christoffel ", qui font intervenir les g_{ij} et les dérivées du premier ordre. Un vecteur mobile dont la différentielle absolue est nulle, c'est-à-dire qui reste constamment équipollent à lui-même, est dit subir un " transport parallèle ".

On peut alors se proposer d'associer à une catégorie d'objets géométriques un certain espace —espace généralisé— fondé sur un groupe G. Si les invariants cherchés peuvent s'exprimer au moyen de notions de géométrie au sens de Klein, on a le moyen de réunir dans une synthèse les approches apparemment divergentes de Riemann et de F. Klein. La théorie du transport parallèle est cependant par elle-même insuffisante à fournir le principe créateur d'une telle synthèse. La géométrie projective classique, sommet de la classification de Klein, ne peut en effet être envisagée que très artificiellement comme théorie de transport parallèle. C'est une fois encore la théorie des groupes qui va fournir ce principe.

Le point important qui se dégage des recherches des géomètres italiens est celui-ci : toutes les notions et tous les résultats de la géométrie riemannienne peuvent être énoncés en termes euclidiens. Si on ne peut parler de déplacements euclidiens dans tout l'espace, de tels déplacements n'existant pas, pas plus qu'il n'existe de figures égales, on peut parler de translations euclidiennes du point A au point infiniment voisin, de tout vecteur en un vecteur équipollent, et de rotations euclidiennes des voisinages du point A autour de A... Dès lors, les notions et les théorèmes relatifs aux vecteurs d'origine A ont leur source dans le groupe des rotations et translations. En bref, et pour s'exprimer de manière intuitive, le ds^2 de l'espace fait du voisinage de tout point A un petit morceau d'espace euclidien et le transport parallèle permet de raccorder ensemble deux morceaux infiniment voisins de manière à les intégrer à un seul et même espace euclidien.

On peut alors se débarrasser de la condition de Levi-Civita, en imposant au transport une condition plus générale liée au groupe.

Tel est le point de départ de la grande généralisation des " espaces à connexion ".

Si l'on exige simplement par exemple du déplacement parallèle qu'il conserve la longueur des vecteurs, on obtient la connexion euclidienne. Un espace à connexion euclidienne n'est rien d'autre qu'une collection de petits morceaux d'espaces euclidiens reliés entre eux de proche en proche, et notamment orientés de proche en proche les uns par rapport aux autres. On définit le parallélisme de deux tangentes en deux points infiniment voisins à une variété V d'un espace euclidien par le fait qu'elles font le même angle avec toute direction tangente à la variété en ce point. La variété est un espace à connexion euclidienne si l'angle de deux tangentes en A est égal à l'angle des deux tangentes en A'. Si le parallélisme se conserve par déformation de la variété, la connexion est " intrinsèque ". On peut ainsi fonder de manière intrinsèque une géométrie euclidienne sur la donnée d'une expression analytique arbitraire de la distance de deux points infiniment voisins, et en particulier sur une forme différentielle quadratique du type $ds^2 = \Sigma\omega_i^2$ L'espace riemannien à ds^2 ainsi obtenu apparaît comme un cas très particulier des espaces à connexion euclidienne, puisqu'à un espace à ds^2 correspond une infinité d'espaces à connexion euclidienne, dont l'un est singulier : il se distingue par la propriété d'avoir une torsion nulle, c'est l'espace euclidien. En retour, la connexion euclidienne détermine complètement le ds^2, qui n'en est qu'un élément particulier [16].

A un espace à connexion euclidienne, on peut associer un sous-groupe du groupe des déplacements euclidiens. Il joue, relativement à l'espace considéré, un rôle analogue au groupe de Galois relativement à l'équation algébrique : de même que ce dernier mesure en quelque sorte l'irrationalité des racines de l'équation, leur écart par rapport à la rationalité, le premier mesure le degré dont l'espace diffère de l'espace euclidien.

On voit facilement que, d'une connexion euclidienne, on peut passer à une connexion affine, une connexion projective, etc.

Si on substitue au groupe des déplacements un " groupe de Lie " transitif quelconque G, on obtiendra les espaces généralisés de groupe $G\ldots$: les points de cet espace généralisé sont les points sur lesquels opère le groupe G.

Ainsi, en général, à tout groupe continu correspondent une *infinité* de géométries, dont celle des espaces riemanniens, qui correspond au groupe des déplacements euclidiens, et qui n'est pas la plus générale —étant de torsion nulle. Mais ces géométries sont non holonomes : dans l'ensemble des géométries correspondant au groupe continu, une seule est holonome (Cartan).

Elie Cartan montre que toute géométrie de Klein, pourvu qu'elle soit à groupe fondamental simple, devient riemannienne par un choix convenable de l'élément générateur de l'espace[17]. Toute géométrie au sens de Klein peut être réalisée comme géométrie rieman-

[16] On retrouve, mais seulement dans un cas assez particulier, le principe d'adjonction / subordination de Klein : il ne s'applique que pour les sous-groupes qui admettent le groupe d'holonomie comme sous-groupe. Voir l'exposé de l'Encyclopédie française [10] (1937), ch. II, La géométrie et la théorie des groupes, Les espaces généralisés et la théorie des groupes, 1.90-7.

[17] Le choix est unique si le groupe est non-unitaire, multiple dans le cas contraire. Cartan cite l'exemple des espaces qu'il appelle " E ", espaces à ds^2 positif dont la courbure riemannienne est conservée par parallélisme, qui ne sont rien d'autre que des espaces de Klein dont le groupe fondamental est le groupe de leurs déplacements. C'est aussi le cas des " espaces symétriques " caractérisée par une des deux propriétés

nienne symétrique par choix de " l'élément générateur ".

Ces espaces sont effectivement plus généraux que les espaces riemanniens, ils jouent par rapport aux espaces kleinéens, le rôle des espaces riemanniens vis-à-vis des espaces euclidiens :

$$\frac{\text{espace généralisé de Cartan}}{\text{espace de Klein}} = \frac{\text{espace riemannien}}{\text{espace euclidien}}$$

Le point de vue très général de Klein laissait de côté la géométrie riemannienne. Après la " vaste synthèse " réalisée par Cartan, cette dernière rentre dans le cadre de la classification kleinéenne, tout en y occupant une place privilégiée

3.3 Signification historique de la généralisation d'E. Cartan

Il n'y avait pas de place pour le schéma de Klein dans la géométrie de Riemann.

Selon le point de vue de Klein, une géométrie est un espace, ou un ensemble de points, avec une structure, et les applications bijectives de l'espace sur lui-même qui conservent la structure forment un groupe, qu'on peut appeler groupe des automorphismes.

Selon Riemann, un automorphisme d'un espace riemannien est une application de l'espace sur lui-même conservant la distance, et il peut arriver que le seul automorphisme de l'espace soit l'identité. D'après Riemann et Helmholtz, seuls les espaces de courbure constante peuvent avoir " suffisamment " d'automorphismes. Aussi la géométrie elliptique, une géométrie à courbure constante positive, a été pendant longtemps appelée " géométrie riemannienne ", de même que la géométrie hyperbolique a été appelée " lobatchevskienne " (ou " bolyaienne " ou " gaussienne ").

D'abord élaboré comme un principe de classification des géométries existantes, et une aide, le programme de Klein devient un dogme : la géométrie est *définie* comme théorie des invariants d'un groupe. Klein, dont l'intérêt pour la physique était grand, pouvait facilement incorporer la théorie de la relativité de 1905 à son schéma : le groupe de Lorentz et sa théorie invariante peuvent être considérées comme une description adéquate de l'électromagnétisme. Mais survient la théorie de la relativité générale, dont la justification physique est pour lui évidente. Klein réalise qu'il doit alors accepter la géométrie riemannienne. Comment faire pour l'incorporer, l'adapter, à la théorie invariante des groupes ? Dans un passage de ses conférences de 1927, il essaie d'annexer au programme d'Erlangen la géométrie riemannienne, en jouant sur l'ambiguïté de la notion de " transformation " : application, ou transformation en coordonnées.

En fait, comme le remarque D. Laugwitz[18], Riemann travaillait à l'intérieur d'une tradition géométrique plus ancienne que celle de Klein et du programme d'Erlangen. Il n'y a pas de coordonnées, ni de mouvements, chez Euclide, et il n'y en a pas dans la formulation moderne des fondements de la géométrie chez Hilbert. On ne vérifie pas la congruence des

suivantes : le tenseur de courbure dérivé est nul, la symétrie conserve les longueurs.

[18]Op. cit. note 11, pp. 249-250. Le texte de Klein se trouve dans un chapitre de la deuxième partie de ses conférences sur " les concepts fondamentaux de théorie invariante et leur pénétration dans la physique mathématique " [11], sur " les groupes de transformations ponctuelles analytiques ". Klein utilise une astuce pour faire en sorte que le programme d'Erlangen semble s'adapter à la géométrie riemannienne. L'argument porte sur l'ambiguïté du terme " transformation ", lequel peut désigner deux choses : soit une application d'un ensemble, sur lui-même ou sur un autre, soit une transformation en coordonnées. Alors que le sens originel était le premier, Klein semble maintenant lui donner le deuxième.

triangles en se donnant la peine de mouvoir le plan entier —une procédure inconcevable pour Euclide . Ce dont on a besoin, c'est de comparaisons de distances. On a vu Riemann introduire les mesures de longueur sur les courbes. En fait, Riemann procède comme les géomètres après Descartes, il calcule en termes de coordonnées spécialement adaptées, normales, inévitablement dérivées de la géométrie. L'existence d'un groupe de mouvements conservant la congruence, dans les cas d'Euclide et de Hilbert, est une conséquence quasi accidentelle, *a posteriori*. C'est avec Klein que coordonnées et applications, qui jouent un rôle secondaire en géométrie, deviennent la définition primordiale de son essence. Or, il est légitime de se demander, d'un point de vue épistémologique, où se trouve, pour la géométrie elle-même, l'objectif souhaitable : ne se trouve-t-il pas, plutôt que dans la théorie invariante d'un groupe de transformations, dans le développement " sans coordonnées " de la géométrie ? Le mathématicien s'y engage à l'absolu, ou à l'intrinsèque, il cherche des invariants.

L'introduction, au cœur de la géométrie, au sens de Riemann, du concept de groupe, qui en est théoriquement absent, permet de déployer toute la richesse cachée de la géométrie, au sens de Klein. Loin de représenter seulement une alternative aux géomètres déconcertés par la réduction de la géométrie à l'algèbre des groupes, elle vient l'enrichir et permet de la généraliser encore. C'est ce qui justifie l'affirmation d'E. Cartan que c'est au fond " sous leur forme riemannienne que les géométries de Klein montre le mieux leurs propriétés fondamentales ". Tout se passe comme si la géométrie riemannienne permettait d'exprimer les virtualités latentes de la géométrie des groupes. L'algébrisation réalisée par Klein risquait de faire perdre de vue l'essence géométrique des problèmes : les idées riemanniennes permettent leur re-géométrisation.

4 Géométrisation de la physique et réflexivité

.

La mathématique, sous la forme que nous lui connaissons aujourd'hui, est fondamentalement réflexive au sens où elle ne cesse d'engendrer et de systématiser, à l'intérieur du corps de connaissance qu'elle constitue, des thèses qui portent sur elle-même en tant que connaissance. Ces thèses constituent la réponse à des problèmes qui ne sont plus exactement des problèmes, mais plutôt des méta-problèmes, des problèmes du second ordre, des problèmes qui sont *sur*, et non plus *de* ou *dans*, la mathématique. Ces méta-problèmes peuvent concerner l'interrogation sur les problèmes à traiter, avec l'identification des problèmes ouverts, significatifs voire profonds, et la manière de les traiter : la pertinence relative des méthodes, le mode de définition des entités, la conduite de la preuve, la nature et le statut des démonstrations, l'organisation de la mathématique en théorie, sa réduction en système, etc. En bref, ils concernent plutôt la justification possible, de la pratique, que la pratique proprement dite, entendue comme conquête de résultats.

On pourrait aller jusqu'à proposer de faire de la réflexivité le trait distinctif des mathématiques, susceptible de les individualiser dans l'ensemble des autres sciences, en tout cas relativement à la physique. Dans aucune autre science, on ne trouve en effet, semble-t-il, ce niveau d'une connaissance proprement réflexive dans lequel une discipline traite les méta- énoncés comme les énoncés, les énoncés qui portent *sur* la discipline en tant que discipline avec les mêmes

outils et les mêmes méthodes que les énoncés *de* la discipline elle-même. Il n'y a pas de séparation entre la discipline et sa matière, son thème ou sujet. C'est une banalité, mais instructive, de faire remarquer qu'en physique, on distingue entre la chaleur et les lois de la chaleur ; alors qu'en mathématiques, les lois, donc la théorie, des objets ou de leurs relations sont elles-mêmes objet. Ainsi les lois de certains êtres, envisagées en elles-mêmes, donnent-elles un nouvel être qui est le groupe, ou un espace, ou une topologie, ou tout cela à la fois, car je peux considérer aussi les objets d'un groupe comme liés par des lois de structure topologique, etc.[19] On pourrait adopter ce trait comme critère pour la distinguer de la physique, laquelle ne peut de la même manière, en tant que discipline, se constituer comme thème de la recherche physique, puisque l'objet de la recherche physique est, jusqu'à preuve du contraire, le monde extérieur, lequel n'inclut pas, sauf thèse réaliste extrême, de partie d'elle-même qui soit constituée de " concept " ou de " théorie ".

La situation est évidemment rendue plus complexe par la mathématisation de la théorie physique, qui rend difficile une séparation nette. Mais on peut se demander si sa forme la plus caractéristique n'est pas constituée par l'importation en physique de la réflexivité mathématique. C'est le phénomène qui nous paraît le plus propre à éclairer la position épistémologique du problème posé par la géométrisation de la théorie physique au sens de Weyl et de Cartan.

La théorie de la relativité a installé au centre de la pensée physique cette véritable catégorie de l'expérience, au sens de Kant, qu'est l'invariance, dont la loi physique, dès lors qu'on lui a procuré son cadre géométrique, constitue l'explicitation des conditions. Cette invariance catégoriale, dont nous avons constaté le pouvoir de stimulation, nous semble constituer en elle-même une remarquable application de la réflexivité mathématique.

Un principe d'invariance apparaît en effet comme un énoncé typiquement réflexif, au sens où il a un statut de méta-énoncé : il porte sur un champ qui est un champ d'énoncés physiques [12]. Si la loi physique est exprimée par une certaine proposition dans une théorie physique (langage), le principe d'invariance consiste à énoncer qu'une certaine autre proposition, exprimant l'invariance de la première, est démontrable dans la théorie physique en question (normalement, et formellement, à partir de ses axiomes). On peut se demander si ce n'est pas l'introduction préalable dans la physique, d'une forme de réflexivité, qui fonde la possibilité de sa géométrisation.

Reprenons, pour préciser ce point, l'exemple du travail d'Hermann Weyl et d'Elie Cartan sur la théorie de la relativité, qui constitue, à cet égard, un véritable point de départ. Un des résultats de Cartan , en parfaite continuité avec les travaux que nous avons évoqués, est qu'un énoncé physique comme celui d'Einstein :

pour un observateur entraîné dans un champ de gravitation et portant avec lui son système de référence en mouvement rectiligne uniforme, les lois de la physique seraient vraies dans son voisinage comme si son système de référence était immobile et comme s'il n'y avait pas de champ de gravitation : son propre mouvement satisferait au principe d'inertie

peut *se traduire* par un énoncé géométrique :

l'espace-temps est une variété à connexion affine.

Une fois exprimée dans le nouveau langage de la géométrie, l'idée d'Einstein revient

[19] C'est le spectacle, familier dans la mathématique d'aujourd'hui, qui, pour certains, est sans doute déjà d'hier, des emboîtements d'ensembles, hiérarchies des structures, superpositions des catégories —jusqu'au sommet où l'on trouve l'ensemble des ensembles, la catégorie des catégories. . .

à considérer l'espace-temps comme une variété à connexion affine. Une telle réduction de la gravitation newtonienne à la géométrie (avec conservation de toutes les notions de la mécanique classique), s'opère par explicitation de la condition d'invariance constitutive du phénomène physique. La géométrie est le moyen même de cette explicitation, et, comme le dit Cartan , on peut alors considérer que se trouve véritablement *expliquée* la gravitation newtonienne : " pour se représenter le mouvement d'un point matériel dans un champ de gravitation comme obéissant à la loi d'inertie, il suffit de considérer l'espace-temps comme un espace non holonome à quatre dimensions dont le groupe fondamental est le groupe de Galilée (avec temps absolu) "[13]. La loi de Newton est alors rigoureusement vraie, son explication se trouve dans la courbure de l'univers : un espace-temps sans courbure correspondrait au champ de gravitation nul ou uniforme, mais ce dernier, ne pouvant être décelé par une expérience mécanique quelconque est identique à un champ nul. La géométrie a permis de rendre compte du phénomène physique. En outre, une fois que l'on se situe dans ce cadre géométrique, et dès lors qu'on donne aux lois de la gravitation newtonienne leur forme invariante —indépendante du système de référence, variable, adopté en différents points de l'univers, le passage à la relativité d'Einstein devient quasiment trivial. Il ne s'agit plus que de faire intervenir la courbure de l'espace-temps, et de substituer, au groupe de Galilée, le groupe de Lorentz. On voit alors que la forme invariante d'Einstein n'est rien d'autre que la transposition, de l'espace-temps de la mécanique classique à celui de la relativité restreinte, de la forme newtonienne invariante de la mécanique classique.

Une fois commencé, ce travail de la remontée réflexive n'a pas véritablement de terme. La géométrie, qui a permis l'expression de l'invariance physique, déploie d'elle-même le champ de ses possibilités et contraintes, induisant avec l'expérience un mouvement de va et vient caractéristique de la physique théorique contemporaine.

Ainsi l'énoncé de H. Weyl qui fonde ses considérations sur la " nature pythagoricienne " de la métrique de l'univers :

pour toute variété métrique de groupe G, il existe une et une seule connexion linéaire sans torsion compatible avec la métrique, apparemment de nature purement mathématique, a également une signification physique. Là encore, il s'agit moins d'un énoncé que d'un méta-énoncé, qui affirme quelque chose sur une théorie, la théorie de certaines structures paramétrées par G. Une part du travail mathématique consiste ici à déterminer une nouvelle axiomatique, susceptible de nous permettre de distinguer, à l'intérieur d'une théorie donnée, la famille des sous-théories pour lesquelles l'énoncé est vrai. On fait alors monter l'axiomatique d'un cran dans une échelle d'abstraction, et cette montée dans l'abstraction appelle une contre-partie intuitive, avec recours à l'expérience. Pour retrouver la théorie initiale, on a besoin de " principes guides "[20]. Tel est le rôle médiateur des principes d'invariance, qui témoignent, au cœur même de la théorie physique, de la présence active de la réflexivité mathématique.

[20] E. Scheibe, *Hermann Weyl and the nature of spacetime*, Exakte Wissenschaften und ihre philosophische Grundlegen, Vorträge des Internationalen Hermann-Weyl Kongresses, ed. W. Deppert et al., Frankfurt am Main, 1988, pp. 73-sq.

Bibliographie

[1] H. Weyl, *Das Raumproblem*, 1921, 1922 ; Ges. Abh., II, pp. 212-228, pp. 328-344 ; *Mathematische Analyse des Raumproblems*, Berlin, 1923

[2] R . Carnap, *Der Raum*, Berlin, 1922 ; *Der Logische Aufbau der Welt*, 1928

[3] G. Granger (*Sur l'idée de concept mathématique naturel*, 1988, repris dans *Formes, opérations, objets*, Paris, 1994, p. 161.

[4] G. Granger, *La pensée de l'espace*, Paris, 1999.

[5] J. Vuillemin, *Physique et métaphysique kantiennes*

[6] Gauss, *Beiträge zur Theorie der algebraischen Gleichungen,* Werke, III : cité par D. Laugwitz, *Bernhard Riemann, 1826-1866, Turning points in the philosophy of mathematics*, Boston-Berlin, 1999, p. 225.

[7] E. Cartan, Encyclopédie française, 1937, ch. II, La géométrie et la théorie des groupes, 1. 88-14.

[8] E. Cartan, *La théorie des groupes et la géométrie*, L'Enseignement mathématique, t. 26, 1927 ; Œuvres d'Elie Cartan, Paris 1952-55, partie I, vol. 2, pp. 841-866.

[9] E. Cartan, surtout *La méthode du repère mobile, la théorie des groupes continus et les espaces généralisés*, 1935, Paris ; Œuvres. . . , op. cit., partie III, vol. 2, pp. 1259-1320.

[10] Encyclopédie française, ch. II, La géométrie et la théorie des groupes, Les espaces généralisés et la théorie des groupes, 1.90-7, 1937.

[11] E. Klein, *Die Grundbegriffe der Invariantentheorie und ihr Eindringenin die Mathematische Physik*, 1927.

[12] E. Wigner, *Symmetries and reflections*, Woodbridge,1979, p. 16-sq.

[13] E. Cartan, *Notice sur les travaux scientifiques* (1931) ; Œuvres. . . , op. cit., partie I, vol . 1, pp. 95-96.

Topologie et cosmologie

Jean-Pierre Luminet
Observatoire de Paris - LUTh
92195 Meudon Cedex

Après Newton les savants surent
Rendre de plus en plus attractifs
Les grands
Et les petits espaces
André Verdet.

Les rapports entre la géométrie, l'astronomie et la cosmologie remontent à l'Antiquité. Certains historiens ont même suggéré que la géométrie s'était développée à partir du VII$^{\text{ème}}$ siècle avant notre ère pour résoudre des problèmes spécifiques de mécanique céleste et d'astronomie sphérique.

La topologie, branche fondamentale de la géométrie, est apparue beaucoup plus tard dans le champ des mathématiques, avec l'*analysis situs* de Leibniz et Euler. Il a fallu attendre deux siècles supplémentaires, c'est-à-dire le XX$^{\text{ème}}$ siècle, pour que la topologie fasse irruption dans le champ de la cosmologie dite relativiste, c'est-à-dire les modèles d'univers fondés sur les équations de la relativité générale.

Malgré cela, *topos* et *cosmos* n'ont guère fait bon ménage dans l'histoire de la cosmologie relativiste – topos et cosmos pouvant être pris au sens étymologique, le *lieu* et la *totalité*, ou bien comme abréviation de deux communautés qui ne se comprennent pas toujours : celles des topologistes et des cosmologistes ! C'est cette histoire que je me propose de résumer, en débordant quelque peu sur les dates butoir 1930-2000 affichées par ce livre. Je commencerai en 1900 et terminerai en 2010 (les astrophysiciens se projettent couramment dans le futur, surtout lorsqu'ils sont en attente de données expérimentales en provenance de projets spatiaux de long terme ... ce qui est précisément le cas pour la topologie cosmique).

> " Imaginez qu'à la suite d'observations astronomiques extrêmement poussées, l'Univers tout entier nous apparaisse constitué d'une multitude de copies conformes de notre propre Voie lactée, que l'espace infini s'avère partagé en cubes contenant chacun une copie exactement identique à la Voie lactée. Pourrait-on réellement croire qu'un même monde se répète indéfiniment ? Pour voir l'absurdité de l'idée, il suffit de songer que dans ces conditions, nous-mêmes, en tant que sujets observateurs, nous existerions à l'identique en un nombre infini d'exemplaires. Il

> serait beaucoup plus agréable de penser que les répétitions sont illusoires, qu'en réalité l'espace possède des propriétés de connectivité particulières, telles que si l'on quitte l'un quelconque de ces cubes par un côté, on y revient immédiatement par le côté opposé " [1]

Cette description, qu'on aura reconnue comme celle d'un espace hypertorique, date d'il y a tout juste un siècle. Elle figure dans le post-scriptum d'un article publié en 1900 par l'astronome allemand Karl Schwarzschild. L'anticipation est d'autant plus extraordinaire que, au début du siècle, les nébuleuses spirales n'avaient pas encore été identifiées comme des galaxies à part entière, extérieures à la nôtre. Karl Schwarzschild faisait partie de ces astronomes visionnaires à la culture mathématique approfondie. Depuis le travail fondateur de Riemann, les mathématiciens du XIX$^{\text{ème}}$ siècle avaient commencé à découvrir des exemples d'espaces de volume fini mais sans frontière. Parmi eux, l'hypertore, constitué d'un seul bloc rectangulaire d'espace euclidien reconnecté sur lui-même, de sorte que tout ce qui passe par une face réapparaît en un point de la face opposée. Mais aux yeux de tous, cet espace clos était une pure abstraction, sans rapport aucun avec l'espace physique. Schwarzschild fut le premier à prendre le contre-pied, à mêler *topos* et *cosmos*, le lieu et la totalité.

Lorsque le père de la relativité, Albert Einstein, introduisit dans son modèle cosmologique de 1917 un espace tridimensionnel de courbure positive, l'hypersphère, sa principale motivation était de fournir un modèle d'espace fini [2]. Cette solution résolvait de façon si ingénieuse tous les paradoxes liés à l'espace newtonien infini que les cosmologistes adhérèrent rapidement à cette idée nouvelle, aux dépens des autres possibilités [22]. Einstein pensait aussi que l'hypersphère fournissait non seulement la métrique de l'espace – c'est-à-dire ses propriétés géométriques locales –, mais aussi sa structure globale, sa topologie ; par exemple, le fait que le volume de l'espace soit fini et égal à $2\pi^2 R^3$, où R est le rayon de l'hypersphère. Or, les questions du volume, de la forme globale de l'espace, et plus généralement de son caractère fini ou infini, ne sont pas seulement affaire de métrique. Elles relèvent avant tout de la topologie. À ce titre, elles exigent une approche supplémentaire par rapport à celle de la géométrie différentielle riemannienne, qui sert de support mathématique à la relativité générale.

La topologie ne faisait pas partie des préoccupations majeures d'Einstein. Dans une lettre à Willem de Sitter, datée du 12 mars 1917, il écrivit en plaisantant : " L'Univers s'étend-il à l'infini ? Est-il fini et clos ? Heine a répondu à cette question dans un de ses poèmes en disant que seuls les imbéciles se la posent. " [4]. Einstein prétendit cependant donner une réponse définitive en faveur de l'Univers clos, avec son modèle d'hypersphère.

Certains de ses collègues, plus au fait des développements de la topologie, lui firent remarquer le caractère arbitraire de son choix. Dès 1917, de Sitter [5] nota que la solution cosmologique d'Einstein admettait une forme différente de l'espace sphérique $\mathbb{S}^3$: l'espace projectif $\mathbb{P}^3$, appelé aussi espace elliptique, construit à partir de l'hypersphère en identifiant deux à deux tous les points situés aux antipodes l'un de l'autre. L'espace elliptique a la même métrique que l'hypersphère, mais une topologie différente ; notamment son volume est moitié moindre. La principale différence est que dans l'hypersphère toutes les géodésiques issues d'un point quelconque reconvergent au point antipodal, tandis que dans l'espace projectif deux géodésiques ont au plus un point commun. Ainsi, la sphère ne satisfait pas au premier axiome d'Euclide , tandis que l'espace projectif y satisfait. Dans $\mathbb{S}^3$ la distance maximale entre deux points quelconques est π, et à partir d'un point donné il n'y a qu'un seul autre

point, l'antipode, situé à cette distance maximale. Dans $\mathbb{P}^3$ la distance maximale est $\pi/2$, et l'ensemble des points situés à la distance maximale forme un plan projectif bidimensionnel $\mathbb{P}^2$. Comme chaque paire de points antipodaux dans l'hypersphère se projette en un point unique dans l'espace projectif, si l'on choisit l'hypersphère pour représenter notre espace, on représente deux fois le monde physique. C'est pour cette raison que de Sitter et, à sa suite, d'autres cosmologistes comme Arthur Eddington [6], argumentèrent que $\mathbf{P}^3$ était réellement le cas le plus simple d'espace clos, et qu'il valait mieux l'adopter dans le contexte cosmologique.

Deux des plus grands mathématiciens de l'époque, H. Weyl et F. Klein, écrivirent également à Einstein pour lui signaler la question topologique. Expert suprême en géométrie et en théorie des groupes, Weyl tenait la chaire de mathématiques à l'Institut polytechnique de Zurich – où Einstein avait effectué de médiocres études. " Mon travail a toujours tenté d'unir la vérité et la beauté ; mais lorsque j'ai dû faire un choix en faveur de l'une ou de l'autre, j'ai toujours choisi la beauté ", écrira-t-il plus tard. Quant à Félix Klein, nul n'était plus habilité que lui à discuter de la topologie de l'espace : en 1872, âgé seulement de 23 ans, il avait présenté pour la rentrée académique de l'université d'Erlangen une dissertation sur la géométrie comprenant en germe un programme d'études destiné à rassembler les disciplines géométriques jusqu'alors dispersées. Moment majeur de l'histoire des mathématiques, à l'origine de l'intégration de la géométrie dans une vue unifiée des mathématiques fondée sur la symétrie, les groupes de transformations et la topologie !

Einstein répondit à Weyl dans une lettre [7] datée de juin 1918 : " J'ai pourtant un sentiment obscur qui me fait préférer le [modèle] sphérique. Je pressens que les variétés dans lesquelles toute courbe fermée peut être contractée continûment en un point sont les plus simples. D'autres personnes doivent également avoir ce sentiment, car autrement, en astronomie, on aurait sans doute pris en considération le cas que notre espace puisse aussi être euclidien et fini. L'espace euclidien à deux dimensions aurait alors les propriétés de connexion d'une surface annulaire. C'est un plan euclidien dans lequel tout phénomène est doublement périodique, où des points qui se trouvent sur la même grille périodique sont identiques. Dans l'espace euclidien fini existeraient trois genres de courbes fermées non continûment réductibles en un point. De façon analogue, l'espace elliptique possède, contrairement au sphérique, une classe de courbes non continûment réductibles en un point ; c'est pourquoi il me plaît moins que le sphérique. Peut-on prouver que l'espace elliptique est la seule variante de l'espace sphérique ? Il me semble. "

Einstein réaffirma sa croyance dans une carte postale du 16 avril 1919, adressée cette fois à Félix Klein [8]. Les termes choisis (" sentiment obscur ", " préférer ", " pressentir ", " plaire "...) indiquent clairement qu'Einstein n'avait aucune argumentation physique à faire valoir pour trancher en faveur du caractère monoconnexe de l'espace ; il était guidé par un *préjugé esthétique*. Dans sa réponse à Weyl , Einstein se trompait aussi sur le dernier point : à trois dimensions, en plus de l'espace elliptique existent une infinité de variantes topologiques de l'espace sphérique, toutes fermées, comprenant notamment les espaces dits " lenticulaires " – alors que, pour les surfaces bidimensionnelles de courbure positive, il n'y a effectivement que deux types topologiques distincts, la sphère et le plan elliptique.

Mais cela, personne ne le savait encore. La classification topologique des espaces tridimensionnels commençait à peine d'être entreprise. L'étude des différentes formes d'espace euclidien avait débuté avec des travaux de cristallographie. En 1885, Feodoroff [9] avait classé

les dix-sept groupes de symétrie pour les structures cristallines, mais ce n'est qu'en 1934 que W. Nowacki [10] démontrera comment les groupes cristallographiques donnent les topologies de l'espace euclidien (figure 1). Le cas des espaces de courbure positive avait été posé par F. Klein en 1890 et par W. Killing en 1891, et il était connu du temps d'Einstein sous le nom de problème de Clifford-Klein [11]. Dès 1904, Poincaré découvrit dans un contexte purement topologique un fascinant exemple de sphère multiconnexe [12]. L'espace sphérique de Poincaré a pour polyèdre fondamental un dodécaèdre régulier de volume 120 fois plus petit que celui de l'hypersphère (figure 2). Le problème de Clifford-Klein ne sera résolu de façon exhaustive qu'en 1930 par Seifert et Threllfall [13] (voir aussi [14] et [15] pour une classification moderne fondée sur les quaternions). Quant au problème des espaces hyperboliques, leur classification ne sera abordée sérieusement que dans les années 1970 sous la houlette de W. Thurston [16] aux Etats-Unis et de Fomenko [17] en Union Soviétique.

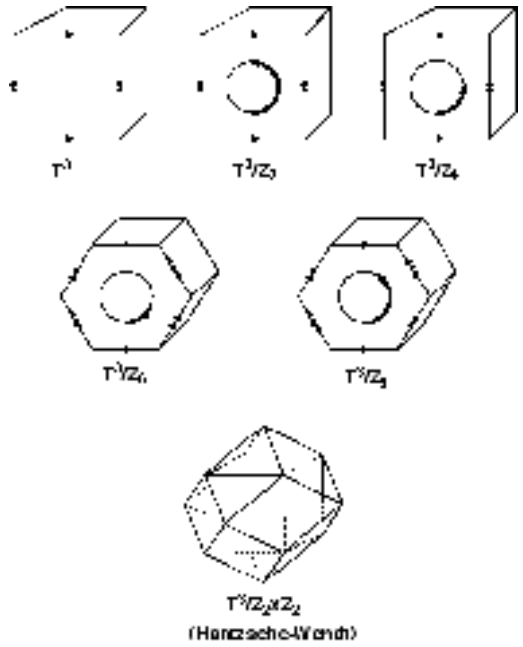


FIG. 8 – **Les six espaces euclidiens fermés orientables**
Parmi les 18 espace euclidiens tridimensionnels, 6 sont de volume fini et orientables. Leurs domaines fondamentaux sont représentés ici, ainsi que leur structure de groupe. $\mathbb{T}^3$ désigne l'hypertore, $\mathbb{Z}_n$ le groupe cyclique d'ordre n.

La découverte de modèles d'univers non statiques, par Friedmann en 1922 [18], puis par Lemaître en 1927 [20], a définitivement changé la face de la cosmologie : l'espace cosmologique a non seulement une courbure, mais aussi une dynamique (voir figure 3). Il est peu connu que leurs articles comportaient également des considérations profondes sur la structure globale de l'Univers. Tous deux prirent en effet rapidement conscience de l'incomplétude de la théorie de la relativité générale sur la question topologique.

La discussion que lui consacra Friedmann se trouve dans l'article de 1924 où il découvrit

FIG. 9 – **L'espace sphérique de Poincaré**
Si chaque face d'un dodécaèdre est collée à sa face opposée après avoir subi une rotation de 36^o (un dixième de tour), l'espace résultant est l'espace elliptique de Poincaré, variante multiconnexe de l'hypersphère, dont le volume est 120 fois plus petit.

les modèles d'univers hyperboliques à expansion perpétuelle [21]. Dans le paragraphe final, il définit clairement les limitations fondamentales de la théorie cosmologique fondée sur la relativité générale :

> " Nous avons démontré que les équations admettent pour solution un univers à courbure spatiale constante négative. Ce fait démontre que les équations d'univers à elles seules ne suffisent pas pour décider si l'univers est fini ou non. Ce que nous savons actuellement de la courbure spatiale ne fournit aucune indication directe sur la finitude ou l'infinitude de l'univers. Des considérations supplémentaires sont nécessaires pour résoudre définitivement la question. En réalité, on peut dire qu'un espace est fini si la distance entre deux points distincts quelconques ne peut pas dépasser une certaine valeur finie. Avant de pouvoir traiter le problème de la finitude de l'espace nous devons donc savoir quels points de l'espace peuvent être considérés comme distincts. [...] Comme critère permettant de distinguer les points entre eux, nous pouvons adopter le principe que l'on peut nommer "Phantomenangst" [la peur des fantômes]. Nous entendons par là l'axiome selon lequel entre deux points distincts on ne peut tracer qu'une seule ligne droite (géodésique). Si nous acceptons ce principe, deux points par lesquels passent plus d'une géodésique ne sont pas distincts. Il en découle par exemple que les deux extrémités d'un diamètre d'une sphère ne devraient pas être considérés comme distincts. Il est clair que ce critère permet l'existence possible de "fantômes", en ce sens qu'en un même point coexistent un objet et ses propres images fantômes. Cette façon de formuler l'identité ou la différence entre les points implique qu'un espace à courbure positive est toujours fini. Elle ne permet cependant pas de résoudre la question de la finitude pour un espace à courbure négative. Ceci justifie notre point de vue selon lequel, en l'absence d'hypothèses additionnelles, les équations d'univers d'Einstein ne permettent pas de trancher la question de la finitude de notre univers. "

Contrairement à Einstein, Friedmann n'avait aucun préjugé en faveur de la topologie monoconnexe, mais il pensait, à l'instar de la plupart des physiciens de l'époque, que seuls les espaces de volume fini étaient physiquement admissibles pour décrire l'espace réel. Les modèles d'univers proposés par Einstein et de Sitter en 1917, et par Friedmann en 1922, avaient une courbure positive et satisfaisaient donc au critère de finitude. Avec les modèles à

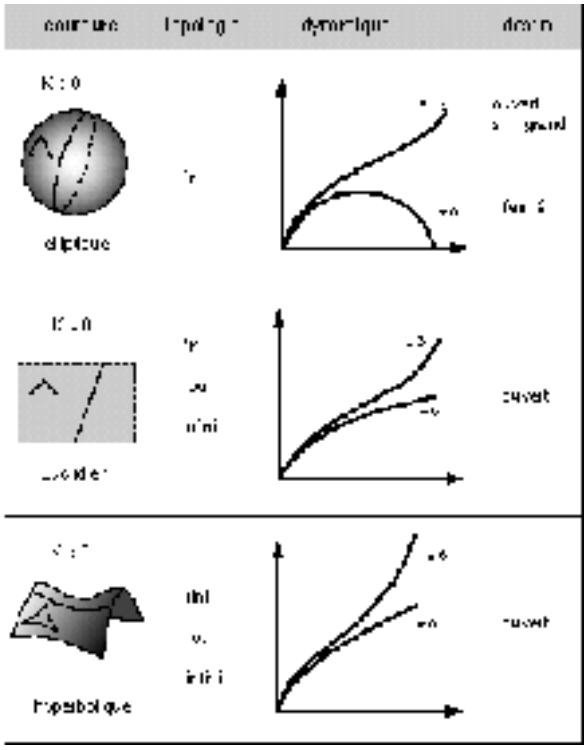


FIG. 10 – **Les solutions cosmologiques de Friedmann – Lemaître.**
Contrairement à une opinion couramment répandue, la courbure de l'espace ne dicte ni l'évolution temporelle de l'Univers (sauf si la constante cosmologique λ est nulle), ni dicte le caractère fini ou infini de l'espace (sauf si la topologie est monoconnexe). Le tableau résume les divers cas pour les solutions de Friedmann – Lemaître, plus connues sour le nom de " modèles de big-bang ". La première colonne indique le signe de la courbure spatiale, la deuxième colonne le caractère fini ou infini de l'espace, la troisième colonne le comportement temporel du facteur d'échelle spatiale, la quatrième colonne le caractère ouvert ou fermé du modèle. Les mesures actuelles tendent à favoriser un modèle à expansion accélérée et à courbure positive ou nulle.

courbure négative, le problème était plus ardu : l'espace hyperbolique, dans sa version la plus simple (topologie monoconnexe), s'étend à l'infini. Donc, même au moment de la " création " de l'Univers, l'espace supposé hyperbolique était lui-même infini ; autrement dit, l'Univers ne serait pas né en un point, mais l'expansion se serait déclenchée en chaque point d'un espace préexistant infini. Conscient de cette difficulté, Friedmann vit une échappatoire dans le fait que les équations d'Einstein ne suffisent pas à décider si l'espace est fini ou infini, même si la courbure est négative ou nulle. Il faut faire des hypothèses supplémentaires pour spécifier des conditions aux limites, notamment le fait de savoir si certains points de l'espace sont identifiés entre eux ou non. Tout le problème de la topologie cosmique était ainsi posé, mais Friedmann ne disposait pas de développements mathématiques suffisants pour poursuivre la discussion.

En effet, à l'époque, aucun espace hyperbolique compact n'était connu. Le premier

exemple sera trouvé par Löbell en 1929 [23], et Weber et Seifert en donneront un exemple plus simple en 1933 [24]. Durant plusieurs décennies, peu d'espaces hyperboliques compacts seront connus. La situation changera brusquement au milieu des années 1970 avec le travail de Thurston, qui montrera que la plupart des variétés tridimensionnelles fermées admettent une géométrie hyperbolique. La classification des espaces hyperboliques fait aujourd'hui encore l'objet de recherches intensives. Les espaces hyperboliques compacts sont notamment classés par volumes croissants à partir du plus petit d'entre eux actuellement connu, l'espace de Weeks [25].

Le rapport historique entre topologie et cosmologie contraste donc avec le rapport entre géométrie non-euclidienne et relativité générale : la géométrie non-euclidienne et le calcul tensoriel en espace riemannien étaient bien développés au moment où les physiciens en eurent besoin pour modéliser la gravitation ; en revanche, les pionniers de la cosmologie relativiste désireux de prendre en compte la topologie dans leurs modèles d'univers n'avaient pas grand-chose à glaner du côté des mathématiques.

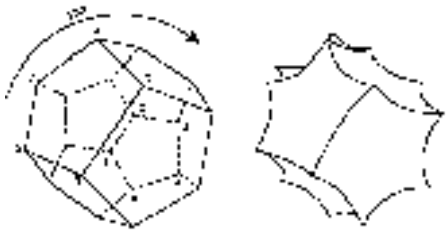


FIG. 11 – :**L'espace hyperbolique de Seifert-Weber**
Si chaque face d'un dodécaèdre (à gauche) est collée à sa face opposée après avoir subi une rotation de 108^o (trois dixièmes de tour), l'espace résultant est l'espace hyperbolique de Seifert-Weber, variante multiconnexe de l'hyperboloïde $\mathbf{H}^3$, mais dont le volume est fini. Le collage ne peut être réalisé que si l'on utilise un dodédaèdre courbé négativement, avec des angles amincis à 72^o (à droite).

Lemaître partageait pleinement le penchant commun en faveur de la finitude de l'espace. Il parla de la géométrie riemannienne comme celle qui avait " dissipé le cauchemar de l'espace infini " [26]. Tous les modèles cosmologiques qu'il construisit à partir de 1927 firent l'hypothèse d'un espace de courbure positive, nécessairement fini, mais ayant la topologie de l'espace projectif et non celle de l'hypersphère. Cela ne l'empêcha nullement de reconnaître d'autres possibilités. Il fut, semble-t-il, le premier cosmologiste à remarquer que les métriques de courbure négative admettaient des topologies de volume fini : " Il est vrai qu'un espace localement hyperbolique n'est pas nécessairement ouvert. Il est possible de construire de tels espaces ayant un volume fini. Cela est même vrai pour l'espace euclidien. " [27]

Malheureusement, les cosmologistes de la première moitié du $XX^{\text{ème}}$ siècle ne disposaient d'aucun moyen expérimental pour mesurer la topologie de l'Univers, de la même façon que Gauss et Lobatchevski, au milieu du $XIX^{\text{ème}}$ siècle, n'avaient aucun moyen de mesurer la courbure de l'espace. Aussi la grande majorité d'entre eux perdit-elle tout intérêt pour la question. Lorsqu'en 1932, Einstein et de Sitter [28] proposèrent le modèle euclidien en expansion perpétuelle, ils firent l'économie de tout ce qui pouvait aller à l'encontre de la simplicité maximale ; ils supposèrent une courbure nulle, une constante cosmologique nulle, et ne prirent même pas la peine de préciser que la topologie était monoconnexe.

Or, c'est cette solution très particulière qui fit office de modèle cosmologique " standard " pendant les soixante années qui suivirent. Ces années-là furent donc les âges sombres de la topologie cosmique. La quasi-totalité des manuels de cosmologie, spécialisés ou non, omit de mentionner la question topologique, réduisant le problème du caractère fini ou infini de l'espace à sa seule courbure. Tous les efforts de la cosmologie observationnelle et théorique portèrent sur la détermination du taux d'expansion et de la courbure de l'espace. Les deux exemples les plus illustres furent un manuel spécialisé de Steven Weinberg [29], qui a constitué une sorte de bible pour tous les étudiants de ma génération, et les livres de vulgarisation de George Gamow [30]. Par son talent de pédagogue et d'humoriste, Gamow imprégna l'imagination de deux générations de jeunes lecteurs, au point d'en convertir certains à la profession de cosmologiste. Le revers de la médaille est que, en affirmant à maintes reprises que les espaces euclidiens et hyperboliques étaient nécessairement infinis, Gamow fut en partie responsable du désintérêt de la profession pour la topologie cosmique !

Par bonheur, ce diktat ne fut pas suivi par tous. Certains " rebelles " s'affirmèrent, au premier rang desquels des théoriciens de la cosmologie quantique, pour qui la question de la topologie de l'Univers était tout à fait naturelle. La cosmologie quantique cherche en effet à comprendre le mécanisme, supposé de nature quantique, par lequel l'Univers serait venu à l'existence. Ils tentent donc de prolonger l'histoire de l'Univers, décrite par la relativité générale et les modèles de big-bang, jusqu'à une époque si reculée qu'il est certain que la relativité générale faillit à la tâche et qu'une généralisation quantique devient nécessaire. Dans les années 1960, John Wheeler fut le premier à suggérer que la topologie de l'espace-temps pouvait fluctuer au niveau microscopique. Naquirent alors des scénarios où le big-bang était remplacé par une naissance spontanée de l'Univers à partir de fluctuations quantiques aléatoires. L'ambition ultime de ces nouvelles approches étant de calculer toutes les propriétés de l'Univers ainsi créé et de concilier les propriétés prédites avec la réalité observable, l'une des propriétés que l'on pouvait espérer prédire étant la topologie de l'espace. Personne ne sut mener à son terme un tel programme, mais du moins l'intérêt pour la topologie dans le cadre de la cosmologie classique était-il relancé.

En 1971, le cosmologiste anglais G. Ellis [31] publia un important article faisant le point sur les récents développements mathématiques concernant la classification des espaces et leurs applications possibles à la cosmologie. Dans la foulée, un regain d'intérêt pour le sujet s'ensuivit, sous la conduite de théoriciens comme le Russe D. Sokoloff, le Brésilien H. Fagundes et le Chinois Fang L.-Z. (par ailleurs dissident politique qui dut émigrer aux États-Unis). Un programme observationnel démarra en Union soviétique, sous la direction de V. Schvartsman. Avec le télescope de six mètres de diamètre nouvellement installé à Zelentchouk, dans le Caucase, les " sources fantômes " que Friedmann avait évoquées en 1924, c'est-à-dire les images multiples d'une même galaxie, furent cherchées. Tous les tests échouèrent : aucune image fantôme de notre Voie lactée ou d'un amas de galaxies proche ne fut reconnue. Ce résultat négatif permit de fixer des contraintes sur la taille minimale d'un espace multiconnexe, mais il n'encouragea guère les chercheurs à poursuivre ce type d'investigation. L'intérêt retomba.

Ce n'est qu'en 1995 que notre volumineux article intitulé " Cosmic topology " vit le jour dans la revue américaine *Physics Reports* [32]. Nous y recensions tout ce que les chercheurs avaient publié sur la question depuis 1917 (une cinquantaine d'articles au total) et indiquions de nouvelles pistes de recherches qui nous paraissaient intéressantes, comme

la méthode de " cristallographie cosmique ". Nous escomptions un regain d'intérêt au niveau international pour ce sujet manifestement négligé, et appelions les compétences jointes des mathématiciens, des théoriciens de la cosmologie et des astronomes. Le fait est que l'époque était mûre pour cela : les cosmologistes disposaient enfin d'outils mathématiques et expérimentaux pour tester la topologie de l'Univers. Les sondages profonds du ciel commençaient à délivrer la géographie tridimensionnelle de l'Univers, et le satellite astronomique COBE venait de scruter en détail la structure du rayonnement fossile, permettant de sonder le passé le plus lointain de l'Univers.

Par une curieuse ironie, deux articles américains parurent la même année, allant complètement à l'encontre de nos espoirs : leurs auteurs déclaraient que le concept de " petit univers " était exclu par les données du satellite COBE et qu'il était inutile de travailler davantage sur la question. Notre premier article détaillant la méthode de cristallographie cosmique fut même refusé par l'*Astrophysical Journal*. Les *referees* répondirent abruptement que la question était close par les données négatives de COBE. Après un bref temps de découragement, nous soumîmes notre article à la revue européenne *Astronomy and Astrophysics*, qui l'accepta [33]. Nous nous aperçûmes aussi que l'interprétation des données de COBE faites par nos collègues américains était quelque peu biaisée, leur analyse tirant une conclusion négative générale à partir des seuls cas extrêmement particuliers – exclusivement euclidiens – qu'ils avaient su traiter.

De fait, ce que nous avions espéré se produisit : on vit enfin fleurir un intérêt mondial pour la topologie cosmique, tant du point de vue théorique qu'observationnel. Dans les cinq ans qui ont suivi – ce qui nous amène à aujourd'hui –, le nombre d'articles parus sur la topologie cosmique a dépassé de loin tous ceux publiés durant les quatre-vingts années précédentes sur la cosmologie relativiste [34]. La forme de l'espace est devenue l'un des problèmes les plus excitants de la cosmologie. Plusieurs dizaines de chercheurs s'y consacrent à travers le monde et les premiers colloques exclusivement consacrés au sujet ont été organisés. La carrière de la topologie cosmique ne fait que commencer, et sa courte histoire est résumée dans la table 1.

Si l'espace est chiffonné, il l'est de façon subtile et à très grande échelle, sinon nous aurions déjà identifié des images fantômes de notre propre galaxie ou d'autres structures bien connues. Or, ce n'est pas le cas. Comment détecter la topologie de l'univers ? Plusieurs équipes, à travers le monde, ont développé des méthodes expérimentales pour tester l'hypothèse d'univers chiffonné.

Deux familles de tests observationnels ont été proposées, toutes deux fondées sur l'effet de mirage topologique. La première s'applique aux sources lumineuses " localisées " comme les galaxies, les quasars et les amas de galaxies, réparties en profondeur dans l'espace et dans le temps ; la seconde utilise le rayonnement fossile, cette source diffuse dans tout l'espace émise par le plasma chaud de l'Univers primitif à une époque unique du passé, et qui nous paraît donc située à la même distance.

Chacune de ces deux familles de tests peut elle-même se subdiviser en deux types : une première méthode tente de reconnaître individuellement des images multiples d'un même objet (mêmes galaxies, mêmes amas, etc. pour les sources localisées, mêmes régions du plasma

Cosmologie	**Topologie**
1687 Newton : $\mathbb{R}^3$	
	1885 Feodoroff : groupes de $\mathbb{R}^3$
	1890 Clifford-Klein : groupes de $\mathbb{S}^3$
1900 Schwarzschild : $\mathbb{T}^3$	
	1904 : Poincaré : $\mathbb{S}^3/\mathbb{I}$
1917 Einstein : $\mathbb{S}^3$	
1917 de Sitter : $\mathbb{P}^3$	
1922 Friedmann : $\mathbb{S}^3$	
1924 Friedmann : $\mathbb{H}^3$, $\mathbb{H}^3/\mathbb{G}$	
1927 Lemaître : $\mathbb{S}^3 - \mathbb{P}^3$, $\mathbb{H}^3$	
	1929 Löbell : premier exemple de $\mathbb{H}^3/\mathbb{G}$
1931 Einstein-de Sitter : $\mathbb{R}^3$	
	1932 Threlfall & Seifert : classification de $\mathbb{S}^3/\mathbb{G}$
	1933 Weber & Seifert : deuxième exemple de $\mathbb{H}^3/\mathbb{G}$
	1934 Nowacki : classification de $\mathbb{R}^3/\mathbb{G}$
1935 Robertson-Walker : $\mathbb{S}^3 - \mathbb{R}^3 - \mathbb{H}^3$	
1958 Lemaître : $\mathbb{H}^3/\mathbb{G}$	1960 Wolf : classification des espaces à courbure constante
1971 Ellis : $\mathbb{M}/\mathbb{G}$	
	1978 Thurston, Fomenko, Weeks... $\mathbb{H}^3/\mathbb{G}$
1995 : Topologie cosmique, “ univers chiffonné ”	

TAB. 1 – **Histoires parallèles de la classification topologique des espaces et de leurs applications à la cosmologie.**
Les symboles $\mathbb{R}^3, \mathbb{S}^3, \mathbb{H}^3$ désignent respectivement les espaces monoconnexes euclidien (infini), hypersphérique (fini) et hyperbolique (infini). $\mathbb{P}^3$ désigne l'espace projectif, la plus simple variante multiconnexe de l'hypersphère, et $\mathbb{T}^3$ désigne l'hypertore, la plus simple variante multiconnexe de l'espace euclidien. Les symboles $\mathbb{R}^3/\mathbb{G}$, $\mathbb{S}^3/\mathbb{G}$ et $\mathbb{H}^3/\mathbb{G}$ désignent les variantes multiconnexes des espaces euclidien, sphérique et hyperbolique, obtenues en identifiant des points par les transformations d'un groupe $\mathbb{G}$.

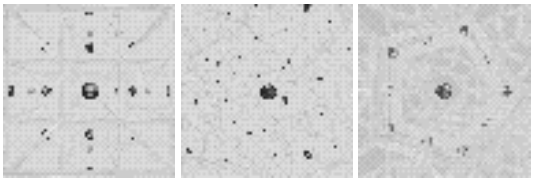


FIG. 12 – Effets de mirage topologique
Si la taille de l'espace est inférieure à celle de l'horizon, l'espace physique, représenté par un domaine fondamental, et les objets cosmiques qu'il contient, se répète un certain nombre de fois dans l'univers observable. Ici, simulations de ces effets de mirage topologique dans un espace euclidien hypertorique (à gauche), un espace hyperbolique de Weeks (au centre) et un espace sphérique de Poincaré (à droite).

émetteur pour le rayonnement diffus) ; l'autre, plus puissante, renonce à la reconnaissance individuelle et s'attache exclusivement à mettre en évidence des corrélations statistiques entre les positions des sources de rayonnement. Pour ce qui est des sources localisées, la cristallographie cosmique [33] tente de repérer certaines répétitions dans la distribution tridimensionnelle des objets lointains ; des histogrammes de séparations de paires sont susceptibles de mettre en évidence un signal topologique sous forme de pics se superposant au bruit de fond produit par une répartition aléatoire des sources. Pour ce qui est du fond diffus cosmologique, la méthode des paires de cercles [35] est fondée sur la distribution angulaire des fluctuations de température du rayonnement fossile. Ce vestige refroidi du big-bang permettrait, si l'espace est chiffonné, de mettre en évidence des corrélations particulières prenant l'aspect de paires de cercles le long desquels les variations de température cosmique d'un point à l'autre seraient les mêmes.

La conclusion provisoire issue de la confrontation de ces méthodes aux données actuellement disponibles est que l'univers ne peut pas être plus petit qu'environ 3 milliards d'années-lumière. Cela laisse encore beaucoup de place pour les modèles d'univers chiffonnés. Compte tenu des dernières contraintes observationnelles sur le paramètre de densité Ω, il reste environ 3000 topologies sphériques candidates pour décrire la structure globale de l'espace – sans compter les 18 formes euclidiennes, au cas improbable où Ω serait strictement égal à 1.

Les projets expérimentaux de cristallographie cosmique et de détection de paires de cercles corrélés sont en cours. Pour l'instant, la profondeur et la résolution des observations ne sont pas suffisantes pour tirer des conclusions sur la topologie globale de l'espace. Mais les prochaines années ouvrent des perspectives enthousiasmantes : elles livreront à la fois des sondages profonds recensant un très grand nombre d'amas lointains de galaxies et de quasars, et des mesures à haute résolution angulaire du rayonnement fossile, grâce aux satellites MAP (2002) et Planck Surveyor (2007). A l'horizon 2010, nous saurons peut-être alors attribuer une forme à l'espace.

Bibliographie

[1] K. Schwarzschild, *Vierteljahrschrift d. Astronom. Gesellschaft*, vol. 35, p. 337 (1900), trad. J.-P. Luminet dans *L'univers chiffonné*, Fayard, 2001.

[2] A. Einstein, *Preussische Akademie der Wissenschaften, Sitzungsberichte*, 1917, p. 142-152.

[3] Pour une discussion complète sur la naissance de la cosmologie relativiste, voir J.-P. Luminet, *L'Invention du big-bang*, suivi de *Essais de cosmologie par A. Friedmann et G. Lemaître*, Le Seuil, coll. Sources du Savoir, 1997.

[4] Archives Einstein, Université de Princeton.

[5] W. de Sitter, *Month. Not. R. Astron. Soc* **78** (1917) 3 ; *Proceedings of the Royal Academy of Amsterdam* **20** (1917) 229.

[6] A. S. Eddington, *The mathematical theory of relativity*, chap. 5, Cambridge University Press, Cambridge, 1923.

[7] Archives Einstein, Université de Princeton, trad. J.-P. Luminet. Voir J.-P. Luminet, *L'univers chiffonné*, op. cit., chap. 44.

[8] Ibid.

[9] E. Feodoroff, *Russian journal for crystallography and mineralogy*, St Petersburg, vol. 21, 1885

[10] W. Novacki, *Commentarii Mathematici Helvetici*, vol. 7, 1934, pp.81-93.

[11] F. Klein, *Mathematisches Annalen*, vol. 37, 1890, p. 544 ; W. Killing, *Mathematisches Annalen*, vol. 39, 1891, p. 257.

[12] H. Poincaré, *Rend. Circ. Mat. Palermo* **18** (1904) 45-110.

[13] W. Threlfall ; H. Seifert, *Math. Annalen* **104** (1930) 1-70, et **107** (1932) 543-586.

[14] J. A. Wolf, *Spaces of constant curvature*, fifth edition, Publish or Perish Inc., Wilmington USA (1967).

[15] E. Gausmann, R. Lehoucq, J.-P. Luminet, J.-P. Uzan, J. Weeks, *Class. Quant. Grav.* **18** (2001) 5155-5186

[16] W. Thurston., *Bull. Am. Math. Soc.* **6** (1982) 357. Une lumineuse synthèse est présentée dans W. Thurston, *Three-Dimensional Geometry and Topology*, Ed. S. Levy, Princeton University Press, 1997. Thurston a reçu la médaille Fields en 1982.

[17] A. Fomenko, *Visual Geometry and Topology*, Springer, 1994.

[18] A. Friedmann, *Zeitschr. für Phys.* **10** (1922) 377-386. Trad. fr. dans

[19]

[20] G. Lemaître, *Ann. Soc. Sci. Bruxelles*, série A, t. XLVII, avril 1927, pp. 29-39 (49-59).

[21] A. Friedmann, *Zeitschr. für Phys.* **2** (1924) 326. Trad. fr. dans

[22]

[23] F. Löbell, *Ber. D. Säches. Akad. D. Wiss.* **83** (1931) 167

[24] C. Weber, H. Seifert, *Math. Zeitschrift* **37** (1933) 237-253.

[25] J. R. Weeks, *PhD thesis*, Princeton University (1985). J.R. Weeks, *SnapPea : A computer program for creating and studying hyperbolic 3-manifolds*, Univ. of Minnesota Geometry Center (en ligne à http ://www.northnet.org/weeks).

[26] G. Lemaître, Texte publié à titre posthume dans " L'Univers, problème accessible à la science humaine ", *Revue d'histoire scientifique*, t. 31, 1978, p. 345-359.

[27] G. Lemaître, *La Structure et l'évolution de l'univers*, 11ème Conseil de physique Solvay, 1958.

[28] A. Einstein, W. de Sitter, *National Academy of Sciences, Proceedings*, vol. XVIII, 1932, p. 213-214. Traduit en français sous le titre dans A. Einstein, *Oeuvres*, *Relativités II*, p. 110-112, Le Seuil.

[29] S. Weinberg, *Gravitation and Cosmology*, New York, J. Wiley, 1972.

[30] G. Gamow, *La Création de l'univers*, Paris, Dunod, 1954 ; *M. Tompkins au pays des merveilles*, Dunod, rééd. 1999.

[31] G. F. R. Ellis, *Gen. Rel. Grav.* **2** (1971) 7.

[32] M. Lachièze-Rey , J.-P. Luminet, *Phys. Rep.* **254** (1995) 135.

[33] R. Lehoucq, M. Lachièze-Rey, J-P. Luminet, *Astron. Astrophys.* **313** (1996) 339.

[34] Voir le volume *Class. Quant. Grav.* **15** (1998) 2529-2732.

[35] N. Cornish, D. Spergel, G. Starkman, *Phys. Rev. Lett.* **77** (1996) 215.

Construction of real algebraic varieties

Ilia Itenberg[1]
CNRS, Institut de Recherche Mathématique de Rennes
Campus de Beaulieu
35042 Rennes Cedex, France

1 Hilbert's 16-th problem

The first part of Hilbert's 16-th problem (see [23]) contains several questions concerning the topology of real algebraic varieties, and more precisely, the topology of plane algebraic curves and spatial algebraic surfaces.

Let us start from the definitions. A *real algebraic curve* of degree m in the real projective plane $\mathbb{R}P^2$ is a real homogeneous polynomial of degree m in three variables considered up to a constant factor. Let A be such a polynomial. Then, the set of points defined by the equation $A(x_0, x_1, x_2) = 0$ in the real projective plane $\mathbb{R}P^2$ is called the *real point set* of the curve A and is denoted by $\mathbb{R}A$. We consider only *nonsingular* curves, which means that corresponding polynomials do not have critical points in $\mathbb{R}^3 \setminus \{0\}$.

In a similar way one defines a *real algebraic hypersurface* of degree m in a real projective space $\mathbb{R}P^n$ as a real homogeneous polynomial of degree m in $n+1$ variables considered up to a constant factor.

The real point set $\mathbb{R}A$ of a nonsingular projective real algebraic curve A is a union of disjoint circles smoothly embedded in $\mathbb{R}P^2$. A circle can be positioned in $\mathbb{R}P^2$ either one-sidedly or two-sidedly. A two-sidedly positioned circle is called an *oval*. An oval divides $\mathbb{R}P^2$ into two parts. The part homeomorphic to a disk is called the *interior* of the oval. All the connected components of the real point set of a nonsingular curve of an even degree in $\mathbb{R}P^2$ are ovals. The real point set of a nonsingular curve of an odd degree in $\mathbb{R}P^2$ contains exactly one one-sidedly positioned connected component. The topological type of the pair $(\mathbb{R}P^2, \mathbb{R}A)$ is defined by the scheme of disposition of the connected components of $\mathbb{R}A$. This scheme is called the *real scheme* of the curve A.

[1] itenberg@maths.univ-rennes1.fr

In the 16-th problem D. Hilbert asked, in particular, to describe all the real schemes of nonsingular curves of a given degree in $\mathbb{R}P^2$. For small degrees (less than 6) the answer is shown in Figures 13, 14 and 15. This answer was known to Hilbert and can be obtained using the Harnack inequality (see below) and the fact that the number of intersection points of a curve of degree m and a line is at most m (if this number of intersection points is finite). The latter statement is a particular case of the Bézout theorem.

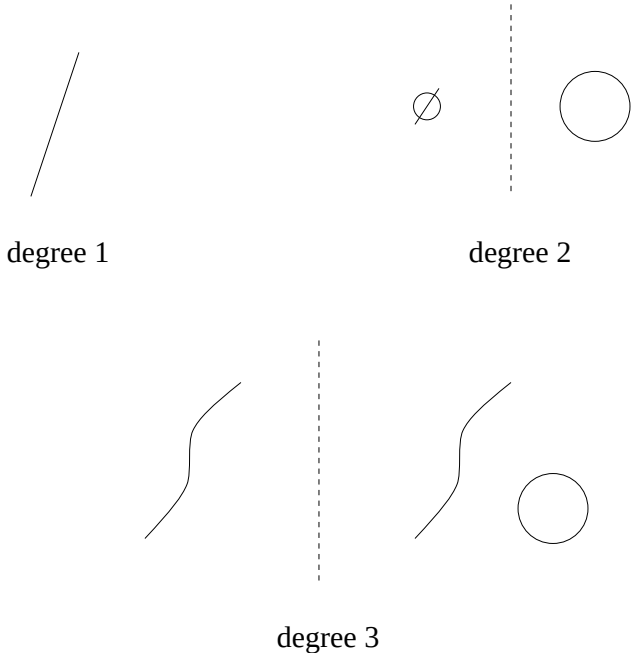


FIG. 13 – Real schemes of nonsingular curves of degrees 1, 2 and 3 in $\mathbb{R}P^2$

Theorem 1.1 (Harnack inequality, see [19]) *The number of connected components of the real point set of a curve of degree m in $\mathbb{R}P^2$ is at most $\frac{(m-1)(m-2)}{2}+1$.*

The curves of degree m in $\mathbb{R}P^2$ whose real point set has $\frac{(m-1)(m-2)}{2}+1$ connected components are called *maximal* (or *M-curves*).

To prove that any real scheme of Figures 13-15 is realizable by a curve of the corresponding degree one can take an appropriate collection of lines and ellipses intersecting transversally and slightly perturb their union (*i.e.*, slightly change the coefficients of the product of the polynomials of degrees 1 and 2 defining these lines and ellipses) in order to obtain the desired picture.

For example, to obtain a maximal curve of degree 4 in $\mathbb{R}P^2$, we can take two ellipses intersecting in 4 points and consider a polynomial $E_1 \cdot E_2 + \varepsilon(x_0^4 + x_1^4 + x_2^4)$, where E_1 and E_2 are polynomials defining two ellipses, and ε is a sufficiently small positive number. If the

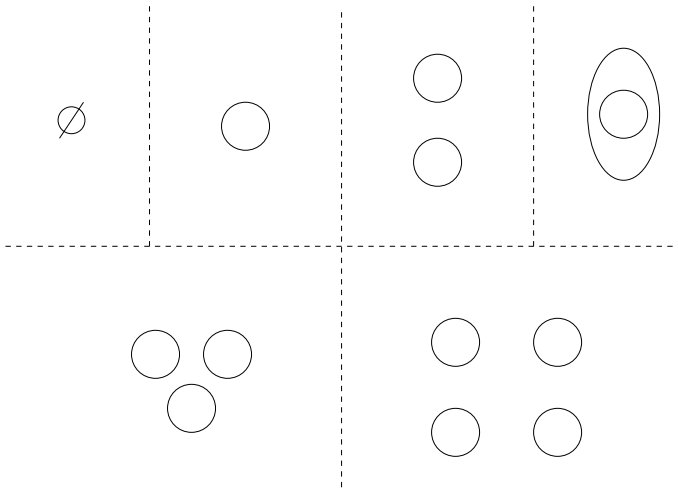


FIG. 14 – Real schemes of nonsingular curves of degree 4 in $\mathbb{R}P^2$

signs of E_1 and E_2 are appropriately chosen, then this polynomial defines a curve whose real point set consists of 4 ovals (see Figure 17).

A classification of real schemes of nonsingular curves of degree 6 in $\mathbb{R}P^2$ is more difficult to obtain. This classification was completed by D. A. Gudkov (see [18]) in 1969. Only 3 real schemes are realizable by maximal curves of degree 6 in $\mathbb{R}P^2$. These schemes are shown in Figure 18.

In 1979 O. Viro [70] obtained a classification of real schemes of nonsingular curves of degree 7 in $\mathbb{R}P^2$. Despite of the efforts to classify the real schemes of nonsingular curves of degree 8 (see, for example, [7, 42, 46, 50, 51, 58, 59, 60, 70, 72]), the classification is still not completed.

A question similar to the first part of Hilbert's 16-th problem concerns a classification of real schemes realized by real pseudo-holomorphic curves of a given degree. A Riemann surface M embedded in $\mathbb{C}P^2$ is a *real pseudo-holomorphic curve*, if it is a J-holomorphic curve in some tame almost complex structure J on X (see [14]) such that $\mathrm{Conj}_* \circ J = J^{-1} \circ \mathrm{Conj}_*$ and $\mathrm{Conj}(M) = M$, where $\mathrm{Conj} : \mathbb{C}P^2 \to \mathbb{C}P^2$ is the involution of complex conjugation. The fixed point set $\mathbb{R}M \subset \mathbb{R}X$ of Conj restricted to a real pseudo-holomorphic curve M is called the *real point set* of M. A real pseudo-holomorphic curve M in $\mathbb{C}P^2$ is *of degree* m if M realizes the class $m[\mathbb{C}P^1] \in H_2(\mathbb{C}P^2)$. It is interesting to compare isotopy classifications of real point sets of real algebraic curves and of real pseudo-holomorphic curves of the same degree. The classifications coincide for the curves of degree ≤ 7. Recently, S. Orevkov [47] obtained an isotopy classification of the real point sets of maximal real pseudo-holomorphic curves of degree 8 (exactly as in the algebraic case, a real pseudo-holomorphic curve of degree m is called *maximal* if its real point set has $\frac{(m-1)(m-2)}{2} + 1$ connected components). As it was already mentioned above, the corresponding classification of real algebraic curves is not yet known.

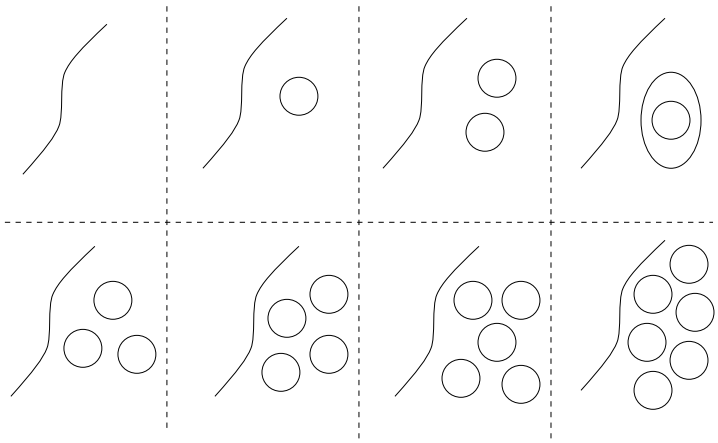


FIG. 15 – Real schemes of nonsingular curves of degree 5 in $\mathbb{R}P^2$

The question on a topological classification (*i.e.*, a classification of the real point sets up to homeomorphism) of nonsingular surfaces of a given degree in $\mathbb{R}P^3$ was also included by Hilbert in the first part of the 16-th problem. At the end of the XIX-th century the answer was known for surfaces of degrees ≤ 3. The topological classification of surfaces of degree 4 in $\mathbb{R}P^3$ was completed by V. Kharlamov [38] in 1976. Kharlamov also obtained finer classifications of surfaces of degree 4 (see [39] and [41]). A topological classification of surfaces of degree ≥ 5 in $\mathbb{R}P^3$ is unknown.

The last three decades of the XX-th century were the years of an intensive progress in topology of real algebraic varieties. Many important restrictions on the topology of real algebraic varieties were obtained using the methods of modern topology of manifolds and complex algebraic geometry (see [1, 54, 55, 56, 34, 35, 36, 37, 11, 44], and also the surveys [16, 77, 2, 74, 10]).

We mention here one general restriction on the topology of real algebraic curves which was obtained in the 1970-s : the Gudkov-Rokhlin congruence. The congruence was conjectured by D. A. Gudkov [15] on the base of the classification of curves of degree 6 and the available examples of curves of higher degrees, and was proved by V. A. Rokhlin [54]. To formulate the congruence, introduce additional definitions and notations. An oval of a nonsingular curve of an even degree in $\mathbb{R}P^2$ is called *even* (resp. *odd*), if it lies inside of even (resp. odd) number of other ovals of the curve. Denote by p (resp., n) the number of even (resp., odd) ovals of a nonsingular curve of an even degree in $\mathbb{R}P^2$.

Theorem 1.2 (Gudkov-Rokhlin congruence, see [54]) *For an M-curve of degree $2k$ in $\mathbb{R}P^2$, on has the congruence*

$$p - n \equiv k^2 \mod 8.$$

Note that the Gudkov-Rokhlin congruence and the Bézout theorem show, in particular, that an M-curve of degree 6 in $\mathbb{R}P^2$ should have one of the real schemes of Figure 18.

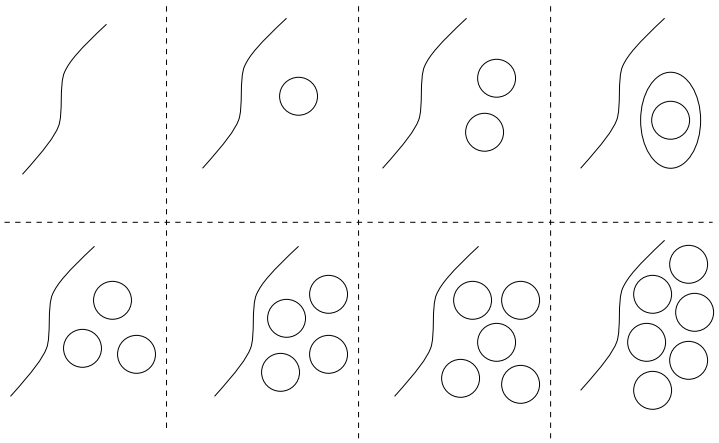


FIG. 16 – Real schemes of nonsingular curves of degree 5 in $\mathbb{R}P^2$

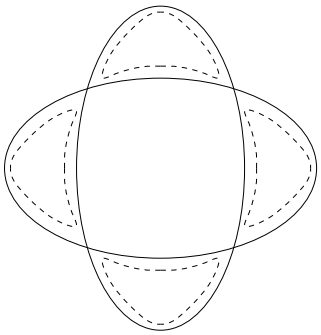


FIG. 17 – Maximal curve of degree 4 in $\mathbb{R}P^2$

The above mentioned progress in topology of real algebraic varieties concerns also the methods of constructions. Until the latest 1970-s, all the constructions of real algebraic curves were based on the method of small perturbations of a product of curves which have smaller degrees and intersect transversally. One can mention here the classical methods of construction proposed by Harnack [19], Hilbert [22], Brusotti [6] and Wiman [78]. To construct several curves of degree 6 in $\mathbb{R}P^2$ Gudkov (see [18]) extended a little bit the classical scheme of construction of curves : he combined the method of small perturbations with the use of quadratic transformations of $\mathbb{R}P^2$.

In 1979, O. Viro proposed a principally new method of construction of real algebraic varieties (see [70, 71, 73, 75, 76] and also [53]). It provides a nice interaction of real algebraic geometry, toric geometry and combinatorics, and gives rise to various generalizations and applications [3, 7, 30, 46, 61, 62, 63, 64, 67]. Almost all the constructions in topology of real algebraic varieties since 1979 use the Viro method.

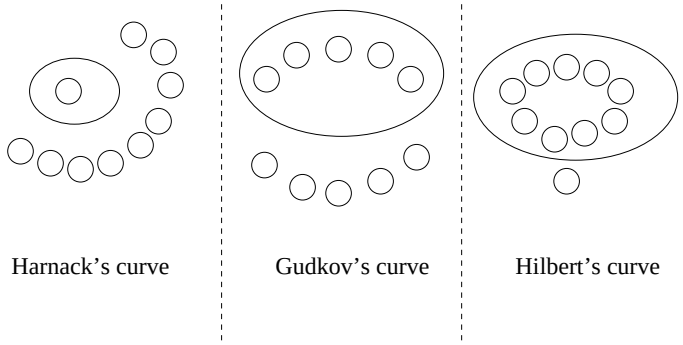


FIG. 18 – Real schemes of maximal curves of degree 6 in $\mathbb{R}P^2$

It is impossible to give here a complete description of the Viro method or a detailed survey of its applications. We restrict ourselves by a discussion of a particular case of the Viro method (the *combinatorial patchworking*) and some of its applications found during the last ten years. The combinatorial patchworking is a powerful construction which gives a possibility to build real algebraic hypersurfaces in a simple combinatorial fashion : one can patchwork them from pieces which essentially are hyperplanes.

2 Combinatorial patchworking

Let m be a positive integer number, and T the triangle in $\mathbb{R}^2$ with vertices $(0,0)$, $(m,0)$ and $(0,m)$. Take a triangulation τ of T with vertices having integer coordinates. Suppose that a distribution of signs at the vertices of τ is given. The sign (plus or minus) at the vertex with coordinates (i,j) is denoted by $\delta_{i,j}$.

Take the copies

$$T_x = s_x(T),\ \ T_y = s_y(T),\ \ T_{xy} = s_x \circ s_y(T)$$

of T, where $s_x,\ s_y$ are reflections with respect to the coordinate axes. Denote by T_* the square $T \cup T_x \cup T_y \cup T_{xy}$. Extend the triangulation τ to a symmetric triangulation of T_*, and the distribution of signs $\delta_{i,j}$ to a distribution at the vertices of the extended triangulation by the following rule : passing from a vertex to its mirror image with respect to a coordinate axis we preserve its sign if the distance from the vertex to the axis is even, and change the sign if the distance is odd.

If a triangle of the triangulation of T_* has vertices of different signs, select a midline joining the middle points of the edges having endpoints of opposite signs. Denote by L the union of the selected segments. It is a piecewise-linear curve contained in T_*. Glue by $s_x \circ s_y$ the opposite sides of T_*. The resulting quotient space $\tilde{T}$ is homeomorphic to $\mathbb{R}P^2$. Denote by $\tilde{L}$ the image of L in $\tilde{T}$.

Let us introduce an additional assumption : the triangulation τ of T is *convex*. This means that there exists a convex piecewise-linear function $\nu : T \longrightarrow \mathbb{R}$ whose domains

of linearity coincide with the triangles of τ. Sometimes, such triangulations are also called coherent (see [13]) or regular (see [79]).

Theorem 2.1 (O. Viro) *If the triangulation τ of T is convex, then there exist a nonsingular real algebraic curve A of degree m in $\mathbb{R}P^2$ and a homeomorphism $\mathbb{R}P^2 \to \tilde{T}$ mapping the set of real points $\mathbb{R}A$ of A onto $\tilde{L}$.*

The described construction is called the *combinatorial patchworking*. An example of a combinatorial patchworking of a curve of degree 3 in $\mathbb{R}P^2$ is shown in Figure 19.

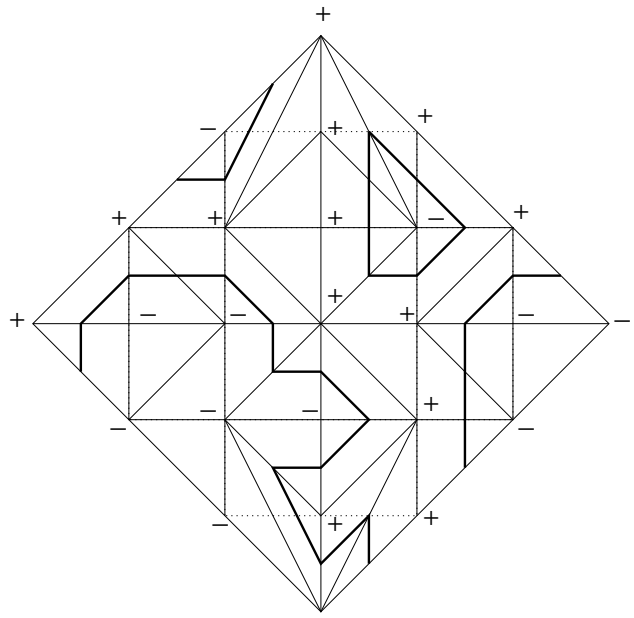


FIG. 19 – Combinatorial patchworking of a curve of degree 3 in $\mathbb{R}P^2$

A curve A obtained by the combinatorial patchworking is called a *T-curve*. A polynomial representing A can be written down explicitly : if t is positive and sufficiently small, the polynomial $\sum_{(i,j)\in V} \delta_{i,j} x_0^i x_1^j x_2^{m-i-j} t^{\nu(i,j)}$ (where V is the set of vertices of τ, and ν is a convex function certifying the convexity of τ) defines a curve with the properties described in Theorem 2.1.

The combinatorial patchworking can be naturally generalized to higher dimensions. The construction starts with the simplex T in $\mathbb{R}^n$ having the vertices $(0,0,0,\ldots,0,0)$, $(m,0,0,\ldots,0,0)$, $(0,m,0,\ldots,0,0)$, …, $(0,0,0,\ldots,0,m)$. One chooses a convex triangulation τ of T such that all the vertices of τ have integer coordinates, and a distribution of signs $\delta : V \to \{+,-\}$ at the set V of vertices of τ. Take the union T_* of 2^n symmetric copies of T under the compositions of reflections with respect to coordinate hyperplanes. Extend τ to a symmetric triangulation τ_* of T_*, and δ to a distribution of signs at the vertices of τ_* using the same rule as above (replacing the words *axis* by the words *hyperplane*). If an n-simplex σ

of τ_* has vertices of different signs, select the convex hull of the middles of edges of σ which have vertices of opposite signs. Let L be the union of all the selected hyperplane pieces, $\tilde{T}$ the quotient space (homeomorphic to $\mathbb{R}P^n$) of T_* under the identification of the points on the boundary which are symmetric with respect to the origin, and $\tilde{L}$ the image of L in $\tilde{T}$.

Theorem 2.2 (O. Viro) *If the triangulation τ of T is convex, then there exist a nonsingular hypersurface A of degree m in $\mathbb{R}P^n$ and a homeomorphism $\mathbb{R}P^n \to \tilde{T}$ mapping the set of real points $\mathbb{R}A$ of A onto $\tilde{L}$.*

Such a hypersurface A in $\mathbb{R}P^n$ is called a *T-hypersurface*.

3 Applications of combinatorial patchworking

3.1 Counterexamples to Ragsdale conjecture

In 1906, V. Ragsdale [52] proposed the following conjecture.

Conjecture 1 (Ragsdale conjecture) *For a nonsingular curve of degree $2k$ in $\mathbb{R}P^2$, one has*

$$p \leq \frac{3k(k-1)}{2} + 1, \quad n \leq \frac{3k(k-1)}{2}.$$

Ragsdale also formulated weaker conjectures. About 30 years later, I. G. Petrovsky [48, 49] proved one of them.

Theorem 3.1 (Petrovsky inequalities, see [49]) *For a nonsingular curve of degree $2k$ in $\mathbb{R}P^2$, one has the inequalities*

$$p - n \leq \frac{3k(k-1)}{2} + 1, \quad n - p \leq \frac{3k(k-1)}{2}.$$

It is clear from [48] and [49] that Petrovsky was not familiar with Ragsdale's paper. He also formulated conjectures on the upper bounds for p and n. His conjecture about n was more cautious (by 1) than Ragsdale's one.

Both the Ragsdale conjecture and Petrovsky's one are wrong. The Ragsdale conjecture for n was disproved by O. Viro [70] in 1979 : in his counterexamples $n = 3k(k-1)/2 + 1$. So, these curves do not contradict to Petrovsky's conjecture.

The first counterexamples to the Ragsdale conjecture for p and to Petrovsky's conjecture for n were constructed in [25] (see also [26] and [32]). In these counterexamples, the difference between p (or n) and $3k(k-1)/2+1$ is a quadratic function of k : they are curves of degree $2k$ in $\mathbb{R}P^2$ with

$$p = \frac{3k(k-1)}{2} + 1 + \left[\frac{(k-3)^2+4}{8}\right],$$

and curves of degree $2k$ in $\mathbb{R}P^2$ with

$$n = \frac{3k(k-1)}{2} + \left[\frac{(k-3)^2+4}{8}\right].$$

These counterexamples were improved by B. Haas [20] who constructed curves of degree $2k$ in $\mathbb{R}P^2$ with

$$p = \frac{3k(k-1)}{2} + 1 + \left[\frac{k^2 - 7k + 16}{6}\right].$$

In the best counterexamples (see [28]) to the Ragsdale conjecture that are currently known, the difference between p and $3k(k-1)/2+1$ (resp., between n and $3k(k-1)/2+1$) is asymptotically equivalent of $9k^2/48$.

The counterexamples to the Ragsdale conjecture are constructed using the combinatorial patchworking, *i.e.*, they are T-curves. The constructions of [25, 26, 32, 28] and [20] are based on the two main observations which are formulated in Lemmas 3.1 and 3.2.

One can also mention here the construction proposed by F. Santos [57]. This construction produces, via the combinatorial patchworking, piecewise-linear curves with more even ovals than in the examples above. However, the triangulation used in Santos' construction is not convex, so one cannot affirm the existence of corresponding algebraic curves. It follows from the results of [31] that the real schemes of Santos' piecewise-linear curves can be realized by real pseudo-holomorphic curves.

A distribution of signs at the integer points of T is called a *Harnack* one if

– any two points (i_1, j_1) and (i_2, j_2) such that

$$i_1 \equiv i_2 \mod 2, \quad j_1 \equiv j_2 \mod 2$$

have the same sign;

– the product of signs at the points $(0,0)$, $(1,0)$, $(0,1)$ and $(1,1)$ is negative.

A triangulation τ of T is called *primitive* if all the triangles of τ are of area $1/2$.

Lemma 3.1 (see [25]) *Let $m = 2k \geq 2$. Then, the combinatorial patchworking applied to an arbitrary primitive convex triangulation of T and a Harnack distribution of signs produces an M-curve with the real scheme shown in Figure 20.*

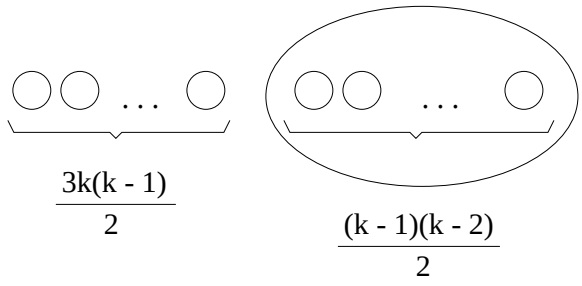


FIG. 20 – Real scheme of the simplest Harnack curve of degree $2k$ in $\mathbb{R}P^2$

Proof. An integer point (i, j) of T is called *interior* if it lies strongly inside T, *even* if both i and j are even, and *odd* if at least one of its coordinates i and j is odd. Permuting the

quadrants and switching if necessary all the signs to the opposite ones (pluses for minuses and minuses for pluses), we can suppose that the distribution of signs is as follows :

all the even points get "−", and all the odd points get "+".

Any even interior point (i, j) of T gives rise to an oval of the curve L. This oval encircles (i, j) and is contained in the star of (i, j). Any odd interior point (i, j) of T also gives rise to an oval of L. This oval encircles the symmetric copy $\sigma(i, j)$ of (i, j) equipped with "+", and is contained in the star of $\sigma(i, j)$. We have found $\frac{(m-1)(m-2)}{2}$ ovals and, thus, the curve L can have only one oval more. This oval does exist, because L intersects the coordinate axes. To finish the proof it remains to notice that the union of the segments

$$\{i - j = -m,\ i \leq 0,\ j \geq 0\} \cup$$
$$\{-m \leq i \leq 0,\ j = 0\} \cup \{i = 0,\ -m \leq j \leq 0\}$$

contains only minuses, and its image in $\tilde{T}$ is not contractible. This means that $3k(k-1)/2$ ovals corresponding to the odd interior points of T (and encircling pluses) are situated outside of the non-empty oval. □

The distribution of signs mentioned in the proof of Lemma 3.1 (all the even points get "−", and all the odd points get "+") is called the *standard Harnack distribution.*

Let S be a diamond equipped with a primitive triangulation and a distribution of signs shown in Figure 21.

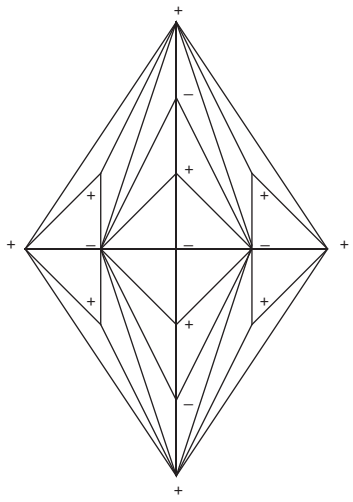


FIG. 21 – Diamond S

Lemma 3.2 (see [25]) *Let $m = 2k$, and suppose that the diamond S is placed inside T in such a way that the center of S has both the coordinates odd. Extend the triangulation of S to a primitive convex triangulation of T. At the integer points of $T \setminus S$ take the standard*

Harnack distribution of signs. Then the combinatorial patchworking applied to the described triangulation and distribution of signs produces a curve with the real scheme shown in Figure 22. The number of even ovals of this curve is equal to $3k(k-1)/2+2$.

Proof. Straightforward. □

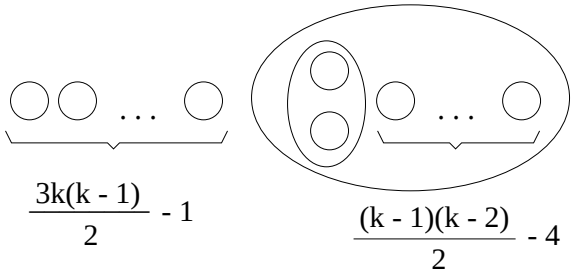


FIG. 22 – Real scheme of a curve of degree $2k$ in $\mathbb{R}P^2$ with $p = 3k(k-1)/2+2$

The combinatorial patchworking of a counterexample to Ragsdale conjecture in degree 10 (for the degrees ≤ 8 the Ragsdale conjecture is true) is shown in Figure 23.

Placing N copies of S in T in such a way that the center of each copy has both the coordinates odd, we can construct a curve of degree $2k$ with $p = 3k(k-1)/2+1+N$. For example, to obtain the counterexamples described in [25], it remains to note that we can place $\left[\frac{(k-3)^2+4}{8}\right]$ copies of S respecting the condition on the coordinates of the centers.

An important open question is to find a sharp upper bound for the number of even (resp., odd) ovals of a curve of degree $2k$ in $\mathbb{R}P^2$. For example, the Harnack and Petrovsky inequalities give the bounds $p \leq 7k^2/4-9k/4+3/2$ and $n \leq 7k^2/4-9k/4+1$, but, probably, these bounds are not sharp. Another question is whether the Ragsdale conjecture is true for M-curves (all the known counterexamples are not M-curves; for maximal T-curves the inequality $p \leq 3k(k-1)/2+1$ was proved by Haas [21]). These questions are directly related to the questions on possible values of Betti numbers of real algebraic surfaces (see Section 3.2).

3.2 Betti numbers of real algebraic surfaces

Let X be a complex algebraic variety, and $c : X \to X$ an antiholomorphic involution. The pair (X, c) is called a *real algebraic variety*. The fixed point set $\mathbb{R}X$ of c is called the *real part* of (X, c). An important question in topology of real algebraic varieties is to determine, for a given complex algebraic variety X, the possible values of Betti numbers of $\mathbb{R}X$. A similar question concerns the possible values of Betti numbers of the real point set of a hypersurface of a given degree in $\mathbb{R}P^n$. For any topological space Y denote by $b_i(Y)$ the dimension of the homology group $H_i(Y;\mathbb{Z}/2)$, and put $b_*(Y) = \sum_i^\infty b_i(Y)$. The first information on the possible values of Betti numbers of $\mathbb{R}X$ is given by the generalized Harnack inequality,

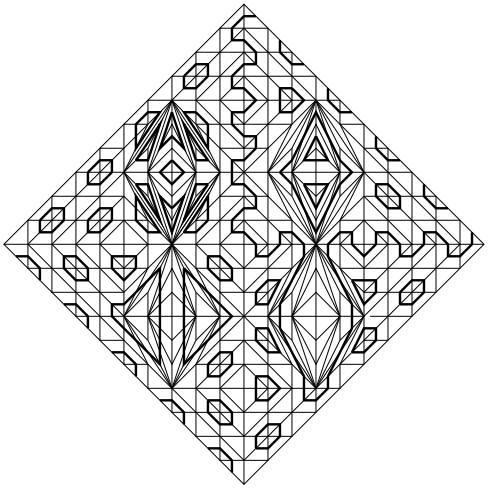


FIG. 23 – Combinatorial patchworking of a counterexample to the Ragsdale conjecture in degree 10 with $p = 32$

generalized Petrovsky-Oleinik inequalities, and Rokhlin and Kharlamov-Gudkov-Krakhnov congruences.

Theorem 3.2 (Generalized Harnack inequality, see [66, 12, 68, 34, 54]) *For a real algebraic variety (X, c), one has the inequality $b_*(\mathbb{R}X) \leq b_*(X)$.*

In the case of real algebraic curves, the inequality of Theorem 3.2 becomes the Harnack inequality. A real algebraic variety (X, c) such that $b_*(\mathbb{R}X) = b_*(X)$ is called *maximal* (or an *M-variety*). Note that for any real algebraic variety (X, c) one has $b_*(\mathbb{R}X) \equiv b_*(X) \mod 2$. If $b_*(\mathbb{R}X) = b_*(X) - 2a$, then the real algebraic variety (X, c) is called an *$(M - a)$-variety.*

Theorem 3.3 (Generalized Petrovsky-Oleinik inequalities, see [36]) *For a compact complex Kähler manifold of dimension $2k$ with an antiholomorphic involution c, one has the inequalities $2 - h^{k,k}(X) \leq \chi(\mathbb{R}X) \leq h^{k,k}(X)$, where $\chi(\mathbb{R}X)$ is the Euler characteristic of the fixed point set $\mathbb{R}X$ of c, and $h^{k,k}(X)$ is the middle Hodge number of X.*

The generalized Petrovsky-Oleinik inequalities incorporate the Petrovsky inequalities (the case of double planes), the Comessatti inequalities (the case of real algebraic surfaces, see [8]) and Petrovsky-Oleinik inequalities (the case of projective hypersurfaces, see [43]).

Theorem 3.4 (Rokhlin congruence, see [54]) *For a maximal real algebraic variety (X, c) of even (complex) dimension, one has the congruence $\chi(\mathbb{R}X) \equiv \sigma(X) \mod 16$, where $\sigma(X)$ is the signature of X.*

The Rokhlin congruence in the case of double planes gives the Gudkov-Rokhlin congruence for real algebraic curves.

Theorem 3.5 (Kharlamov-Gudkov-Krakhnov congruence, see [35] and [17]) *For a real algebraic $(M-1)$-variety (X, c) of even (complex) dimension, one has the congruence $\chi(\mathbb{R}X) \equiv \sigma(X) \pm 2 \mod 16$.*

We give here a brief survey of known results on the Betti numbers of real algebraic surfaces in 3-dimensional projective space (for a detailed information on general real algebraic surfaces see [9, 10, 40, 65, 77]). For real algebraic surfaces the question on the individual Betti numbers of the real part reduces to that on the number of connected components and the number of handles. It turns out that almost all recent results concerning the possible values of Betti numbers of real algebraic surfaces in $\mathbb{R}P^3$ are constructions of surfaces with record values of individual Betti numbers.

Let us start with the discussion of the following question : *what is the possible values of $b_0(\mathbb{R}X)$, where X is a nonsingular surface of degree m in $\mathbb{R}P^3$?* The answer is known for $m \leq 4$, *i.e.*, exactly in the cases when the topological classification is available. It follows from the generalized Harnack and Comessatti inequalities, that the number of connected components of the real point set of a nonsingular surface of degree m in $\mathbb{R}P^3$ is at most $5m^3/12-3m^2/2+25m/12$. In particular, $b_0(\mathbb{R}X) \leq 25$ for a nonsingular surface X of degree 5 in $\mathbb{R}P^3$. No nonsingular surface of degree 5 in $\mathbb{R}P^3$ with 24 or 25 connected components is known. A nonsingular surface of degree 5 with 23 connected components was constructed by S. Orevkov (see [45]). The construction presented in [45] is a modification of the construction of a nonsingular surface of degree 5 with 22 connected components proposed in [29]. For algebraic surfaces of higher degrees in $\mathbb{R}P^3$, the asymptotic difference between the number of connected components in the best examples currently known and the estimate $5m^3/12 - 3m^2/2 + 25m/12$ is big enough. F. Bihan [3] proved that for any given $\epsilon > 0$, there exist a family $\{X_m,\ m \in \mathfrak{N}\}$, where X_m is a nonsingular surface of degree m in $\mathbb{R}P^3$, and a constant C such that

$$b_0(\mathbb{R}X_m) > (\frac{13}{36} - \epsilon) \cdot m^3 - C.$$

Consider now the similar question on the first Betti number : *what is the possible values of $b_1(\mathbb{R}X)$, where X is a nonsingular surface of degree m in $\mathbb{R}P^3$?* Again the answer is known only for $m \leq 4$. The generalized Harnack and Comessatti inequalities imply that the first Betti number of the real point set of a nonsingular surface of degree m in $\mathbb{R}P^3$ is at most $5m^3/6 - 3m^2 + 25m/6 - 1$. In particular, $b_1(\mathbb{R}X) \leq 49$ for a nonsingular surface X of degree 5 in $\mathbb{R}P^3$. A surface of degree 5 with the first Betti number of the real part equal to 49 would be an M-surface and would contradict to the Rokhlin congruence. Hence, for a nonsingular surface X of degree 5 in $\mathbb{R}P^3$, one has $b_1(\mathbb{R}X) \leq 47$.

O. Viro [74] proposed the following conjecture : *let X be a nonsingular simply connected compact complex surface with an antiholomorphic involution $c : X \to X$; then $b_1(\mathbb{R}X) \leq h^{1,1}(X)$, where $\mathbb{R}X$ is the fixed point set of c.* This conjecture would imply that the first Betti number of the real point set of a nonsingular surface of degree m in $\mathbb{R}P^3$ is at most $2m^3/3 - 2m^2 + 7m/3$, and in particular, that $b_1(\mathbb{R}X) \leq 45$ for a nonsingular surface X of degree 5 in $\mathbb{R}P^3$.

Any nonsingular curve of degree $2k$ in $\mathbb{R}P^2$ with $p > \frac{3k^2-3k+2}{2}$ produces a counterexample to Viro's conjecture : one can take a double plane ramified along the complex point set of a counterexample to the Ragsdale conjecture with an appropriate choice of a lifting of the

involution of complex conjugation. Thus, the counterexamples to Ragsdale conjecture obtained in [25] show that Viro's conjecture is not true. Counterexamples to Viro's conjecture among algebraic surfaces in $\mathbb{R}P^3$ where constructed in [27] and [4] (the counterexamples of [27] are of degrees ≥ 10; the counterexample of [4] is of degree 6). Recall that a counterexample in degree 5 would have the first Betti number of the real part equal to 47. Bihan [5] constructed a real numerical quintic surface X with $b_1(\mathbb{R}X) = 47$, and, hence, showed that the maximal possible first Betti number of the real part of a real numerical quintic is 47. It is not known whether a surface of degree 5 in $\mathbb{R}P^3$ can have the first Betti number of the real part equal to 47.

For algebraic surfaces of higher degrees in $\mathbb{R}P^3$, again the asymptotic difference between the first Betti number in the best examples currently known and the estimate $5m^3/6 - 3m^2 + 25m/6 - 1$ is big enough. Bihan [3] proved that for any given $\epsilon > 0$ there exist a family $\{X_m,\ m \in \mathfrak{N}\}$, where X_m is a nonsingular surface of degree m in $\mathbb{R}P^3$, and a constant C such that

$$b_1(\mathbb{R}X_m) > (\frac{13}{18} - \epsilon) \cdot m^3 - C.$$

Note that any improvement of counterexamples to the Ragsdale conjecture gives automatically, via Bihan's construction [3], a new asymptotic lower bounds for the maximal number of connected components and the maximal first Betti number of a surface of a given degree in $\mathbb{R}P^3$.

The constructions of real algebraic surfaces with record Betti numbers mentioned in this section are, in a sense, based on the construction of counterexamples to the Ragsdale conjecture [25, 26, 32]. All these constructions use combinatorial patchworking. Some of them ([29, 45, 5]) combine the combinatorial patchworking with a construction of quintic surfaces due to E. Horikawa (see [24]). Two Bihan's constructions (see [3] and [4]) combine the combinatorial patchworking with the doubling construction (the latter produces surfaces of degree $2k$ in $\mathbb{R}P^3$ given by polynomials of the form $P(x_0, x_1, x_2, x_3) = P_k(x_0, x_1, x_2, x_3)^2 + tP_{2k}(x_0, x_1, x_2, x_3)$, where P_k and P_{2k} are homogeneous polynomials of degrees k and $2k$, respectively, and $t > 0$ is a small parameter). The main fragment used in the record constructions is either the diamond S (see Figure 21) or a half of S (as in the constructions of [29] and [45]).

Another important application of the combinatorial patchworking concerning the Betti numbers of real algebraic varieties is a construction, for any natural m and n, of a maximal hypersurface of degree m in $\mathbb{R}P^n$ (see the forthcoming paper [33]).

Bibliographie

[1] V. I. Arnold, *On arrangement of ovals of real plane algebraic curves, the involutions of four-dimensional smooth manifolds, and the arithmetic of integral quadratic forms.* Funkc. Anal. i Prilozhen. **5** (1971), no. 3, 1-9.

[2] V. I.Arnold, O. A. Oleinik, *Topology of real algebraic manifolds*, Vestnik Mosk. Univ., Ser. I, Mat. i Mekh. **6** (1979), 7-17.

[3] F. Bihan, *Constructions combinatoires de surfaces algébriques réelles.* Ph. D. Thesis. Rennes. 1998.

[4] F. Bihan, *Une sextique de l'espace projectif réel avec un grand nombre d'anses*, Rev. Mat. Complut. **14** (2001), no. 2, 439-461.

[5] F. Bihan, *Betti numbers of real numerical quintic surfaces. In : Topology, ergodic theory, real algebraic geometry*, Amer. Math. Soc. Transl. Ser. 2, **202**, Amer. Math. Soc., Providence, RI, 2001, pp. 31-38.

[6] L. Brusotti, *Sulla generazzione di curve piane algebriche reali mediante "piccola variazione" di una curve spezzata*, Annali di Mat. (3) **22** (1913), 117-169.

[7] B. Chevallier, *Four M-curves of degree* 8, Funct. Anal. Appl. (to appear).

[8] A. Comessatti, *Sulla connessione delle superficie algebriche reali* in *Verhandl. Intern. Math. Kongr. Zürich* **2** (1932), 169.

[9] A. Degtyarev, I. Itenberg, V. Kharlamov, *Real Enriques surfaces.* Lect. Notes in Math. **1746**, Springer-Verlag, Berlin, 2000.

[10] A. Degtyarev, V. Kharlamov, *Topological properties of real algebraic varieties : Rokhlin's way*, Uspekhi Mat. Nauk **55** (2000), no. 4 (334), 129-212 (Russian).

[11] T. Fiedler, *Pencils of lines and the topology of real algebraic curves*, Izv. Akad. Nauk SSSR Ser. Mat. **46** (1982), no.4, 853-863 (Russian) ; English transl. in *Math. USSR-Izv.* **21** (1983), 161-170.

[12] E. E. Floyd, *On periodic maps and the Euler characteristics of associated spaces*, Trans. AMS **72** (1952), 138-147.

[13] I. M. Gelfand, M. M. Kapranov, A. V. Zelevinski, *Discriminants, resultants and multidimensional determinants*, Birkhäuser, Boston, 1994.

[14] M. Gromov, *Pseudo-holomorphic curves in symplectic manifolds*, Invent. Math. **82** (1985), 307-347.

[15] D. A. Gudkov, *Construction of a new series of M-curves*, Dokl. Akad. Nauk SSSR **200**, (1971), no. 6, 1269-1272.

[16] D. A. Gudkov, *Topology of real projective algebraic varieties*, Rus. Math. Surveys **29** (1974), no. 4, 3-79.

[17] D. A. Gudkov, A. D. Krakhnov, *The periodicity of the Euler characteristic of real algebraic* $(M-1)$*-manifolds*, Funkc. Anal. i Prilozhen. **7** (1973), no. 2, 15-19 (Russian).

[18] D. A. Gudkov, G. A, Utkin, *Topology of curves of order 6 and surfaces of order 4* in *Uch. Zapiski Gor′kov. Univ.* **87** (1969).

[19] A. Harnack, *Über Vieltheiligkeit der ebenen algebraischen Curven*, Math. Ann **10** (1876), 189-199.

[20] B. Haas, *Nouveaux contre-exemples à la conjecture de Ragsdale*, C. R. Acad. Sci. Paris, Série I. **320** (1995), 1507-1512.

[21] B. Haas, *The Ragsdale conjecture for maximal T-curves*, Preprint, 1997, Universität Basel.

[22] D. Hilbert, *Über die reelen züge algebraischen Curven*, Math. Ann. **38** (1891), 115-138.

[23] D. Hilbert, *Mathematische probleme*, Arch. Math. Phys. (3) **1** (1901), 213-237.

[24] E. Horikawa, *On deformations of quintic surfaces*, Invent. Math. **31** (1975), 43-85.

[25] I. Itenberg, *Contre-exemples à la conjecture de Ragsdale*, C. R. Acad. Sci. Paris, Série I. **317** (1993), 277-282.

[26] I. Itenberg, *Counterexamples to Ragsdale conjecture and T-curves* in *Contemporary Math.* **182** (Proc. Conf. Real Alg. Geom., December 17-21, 1993, Michigan, ed. S.Akbulut), AMS, Providence, R.I., 1995, pp. 55-72.

[27] I. Itenberg, *Topology of real algebraic T-surfaces*, Revista Mat. de la Univ. Compl. de Madrid (Proceedings, Segovie 1995) **10** (1997), 131 - 152.

[28] I. Itenberg, *On the number of even ovals of a nonsingular curve of even degree in* $\mathbb{R}P^2$ in *Topology, ergodic theory, real algebraic geometry*, Amer. Math. Soc. Transl. Ser. 2, **202**, Amer. Math. Soc., Providence, RI, 2001, pp. 121-129.

[29] I. Itenberg, V. Kharlamov, *Towards the maximal number of components of a nonsingular surface of degree* 5 *in* $\mathbb{R}P^3$, Amer. Math. Soc. Trans.(2). **173** (1996), 111-118.

[30] I. Itenberg, E. Shustin, *Singular points and limit cycles of planar polynomial vector fields*, Duke Math. J. **102** (2000), no. 1, 1-37.

[31] I. Itenberg, E. Shustin, *Combinatorial patchworking of real pseudo-holomorphic curves*, to appear in Proceedings of 8-th Gökova Geometry-Topology Conference, 2001.

[32] I. Itenberg, O. Viro, *Patchworking algebraic curves disproves the Ragsdale conjecture*, Math. Intelligencer **18** (1996), no. 4, 19-28.

[33] I. Itenberg, O. Viro, *Maximal real algebraic hypersurfaces of projective space*, (in preparation).

[34] V. Kharlamov, *The maximal number of components of a 4th degree surface in* $\mathbb{R}P^3$, Funkc. Anal. i Prilozhen. **6** (1972), no. 4, 101.

[35] V. Kharlamov, *New congruences for the Euler characteristic of real algebraic manifolds*, Funkc. Anal. i Prilozhen. **7** (1973), no. 2, 74-78 (Russian) ; English transl. in *Funct. Anal. Appl.* **7** (1973), 147-150.

[36] V. Kharlamov, *A generalized Petrovskii inequality*, Funkc. Anal. i Prilozhen. **8** (1974), no. 2, 50-56 (Russian) ; English transl. in Funct. Anal. Appl. **8** (1974), no. 2, 132-137.

[37] V. Kharlamov, *Additional congruences for the Euler characteristic of even-dimensional real algebraic manifolds*, Funct. Anal. Appl. **9** (1975), no. 2, 134-141.

[38] V. Kharlamov, *The topological type of nonsingular surfaces in* $\mathbb{R}P^3$ *of degree 4*, Funkc. Anal. i Prilozhen. **10** (1976), no. 4, 55-68 (Russian) ; English transl. in Funct. Anal. Appl. **10** (1976), no. 4, 295-305.

[39] V. Kharlamov, *Isotopy types of non-singular surfaces of degree 4 in* $\mathbb{R}P^3$, Funct. Anal. Appl. **12** (1978), 86-87.

[40] V. Kharlamov, *Real algebraic surfaces* in *Proc. Internat. Congress Math., Helsinky* **1** (1978), 421-428 (in Russian).

[41] V. Kharlamov, *On the classification of non-singular surfaces of degree 4 in* $\mathbb{R}P^3$ *with respect to rigid isotopies*, Funct. Anal. Appl. **18** (1984), no. 1, 49-56.

[42] A. Korchagin, *New M-curves of degrees* 8 *and* 9, Dokl. Akad. Nauk SSSR **306** (1989), no. 5, 1038-1041 (Russian) ; English transl. in Soviet Math. Dokl. **39** (1989), no. 3, 569-572.

[43] O. A. Oleinik, I. G. Petrovsky, *On the topology of real algebraic surfaces*, Isv. Akad. Nauk SSSR, Ser. Mat. **13** (1949), 389-402 (Russian).

[44] S. Yu. Orevkov, *Link theory and oval arrangements of real algebraic curves*, Topology **38** (1999), 779-810.

[45] S. Yu. Orevkov, *Real quintic surface with* 23 *components*, C. R. Acad. Sci. Paris, Série. I. **333** (2001), no. 2, 115-118.

[46] S. Yu. Orevkov, *New M-curve of degree 8*, to appear.

[47] S. Yu. Orevkov, *Classification of flexible M-curves of degree 8 up to isotopy*, Geometric and Functional Analysis, to appear.

[48] I. G. Petrovsky, *Sur la topologie des courbes réelles et algébriques*, C. R. Acad. Sci. Paris **197** (1933), 1270-1272.

[49] I. G. Petrovsky, *On the topology of real plane algebraic curves*, Ann. Math. **39** (1938), no. 1, 189-209.

[50] G. M. Polotovsky, *(M - 2)-curves of order* 8 *and some conjectures*, Uspekhi Mat. Nauk **36** (1981), no. 4, 235-236.

[51] G. M. Polotovsky, *On the classification of nonsingular curves of degree* 8, In *Topology and geometry—Rohlin Seminar*, Lecture Notes in Math. **1346**, Springer, Berlin, 1988, pp. 455-485.

[52] V. Ragsdale, *On the arrangement of the real branches of plane algebraic curves*, Amer. J. Math. **28** (1906), 377-404.

[53] J.-J. Risler, *Construction d'hypersurfaces réelles (d'après Viro])*, Séminaire N.Bourbaki, no. 763, vol. 1992-93, Novembre 1992.

[54] Rokhlin, V. A. : *Congruences modulo 16 in Hilbert's 16th problem*, Funct. Anal. Appl. **6** (1972), 301-306.

[55] V. A. Rokhlin, *Complex orientation of real algebraic curves*, Funkc. Anal. i Prilozhen. **8** (1974), no. 4, 71-75.

[56] V. A. Rokhlin, *Complex topological characteristics of real algebraic curves*, Russ. Math. Surveys *33* (1978), no. 5, 85-98.

[57] F. Santos, *Improved counterexamples to the Ragsdale conjecture*, preprint, Universidad de Cantabria, 1994.

[58] E. Shustin, *Independence of smoothings of singular points and new M-curves of degree* 8, Uspekhi Mat. Nauk **40** (1985), no. 5, 212.

[59] E. Shustin, *A new M-curve of degree eight*, Mat. Zametki **42** (1987), no. 2, 180–186 (Russian).

[60] E. Shustin, *New M- and $(M-1)$-curves of degree* 8, in *Topology and geometry—Rohlin Seminar*, Lecture Notes in Math. **1346**, Springer, Berlin, 1988, pp. 487-493.

[61] E. Shustin, *Topology of real plane algebraic curves*, in *Proc. Intern. Conf. Real Algebraic Geometry, Rennes, June 24-29 1991*. Lecture Notes in Math. **1524**, Springer, 1992, pp. 97-109.

[62] E. Shustin, *Real plane algebraic curves with prescribed singularities*, Topology **32** (1993), no. 4, 845-856.

[63] E. Shustin, *Surjectivity of the resultant map : a solution to the inverse Bezout problem*, Comm. Algebra **23** (1995), no. 3, 1145-1163.

[64] E. Shustin, *Gluing of singular and critical points*, Topology **37** (1998), no. 1, 195-217.

[65] R. Silhol, *Real Algebraic Surfaces*, Lecture Notes in Math. **1392**, 1989, Springer-Verlag.

[66] P. A. Smith, *Fixed points of periodic transformations*, Appendix to S. Lefschetz, *Algebraic topology*. Amer. Math. Soc. Colloquium Publ. **27**. American Mathematical Society, New York, 1942.

[67] B. Sturmfels, *Viro's theorem for complete intersections*, Annali della Scuola Normale Superiore di Pisa (4) **21** (1994), no. 3, 377-386.

[68] R. Thom, *Sur l'homologie des variétés algébriques réelles*, in *A symposium in honour of Marston Morse*. Princeton Univ. Press, 1965, pp. 255-265.

[69] O. Ya. Viro, *Construction of M-surfaces*, Funct. Anal. Appl. **13** (1979), no. 3.

[70] O. Ya. Viro, *Curves of degree 7, curves of degree 8 and the Ragsdale conjecture*, Dokl. Akad. Nauk SSSR **254** (1980), no. 6, 1306-1310 (Russian) ; English transl. in Soviet Math. Dokl. **22** (1980), 566-570.

[71] O. Ya. Viro, *Gluing of algebraic hypersurfaces, smoothing of singularities and construction of curves*, in *Proc. Leningrad Int. Topological Conf., Leningrad, Aug. 1982*, Nauka, Leningrad, 1983, pp. 149-197 (Russian).

[72] O. Ya. Viro, *Real plane curves of degrees 7 and 8 : new prohibitions*, Izv. Akad. Nauk SSSR, Ser. Mat. **47** (1983), 1135-1150 (Russian) ; English transl. in Math. USSR Izvestia **23** (1984), 409-422.

[73] O. Ya. Viro, *Gluing of plane real algebraic curves and construction of curves of degrees* 6 *and* 7, in Lect. Notes Math. **1060**, Springer, Berlin etc., 1984, pp. 187-200.

[74] O. Ya. Viro, *Progress in the topology of real algebraic varieties over the last six years*, Rus. Math. Surv. **41** (1986), no. 3, 55-82.

[75] O. Ya. Viro, *Real algebraic plane curves : constructions with controlled topology*, Leningrad Math. J. **1** (1990), 1059-1134.

[76] O. Ya. Viro, *Patchworking real algebraic varieties*, preprint, Uppsala University, 1994.

[77] G. Wilson, *Hilbert's sixteenth problem*, Topology **17** (1978), no. 1, 53-73.

[78] A. Wiman, *Über die reellen Züge der ebenen algebraischen Kurven*, Math. Ann. **90** (1923), 222-228.

[79] G. Ziegler, *Lectures on polytopes*, Graduate Texts in Mathematics **152**, Springer-Verlag, Berlin, 1995.

Some Topologies for Computations

Giuseppe Longo
CNRS and Déptartement d'Informatique
Ecole Normale Supérieure
45, rue d'Ulm
75005 Paris, France
www.di.ens.fr/users/longo

1 Introduction

1.1 Computations

The applications of topological and order structures in Theory of Computation, a key aspect of Foundations of Mathematics and of Theoretical Computer Science, has various origins and it is largely due to the relevance of these structures for Intuitionistic (constructive) Systems : their Logic underlies theoretical as well as applied Computability.

Computability Theory was born in the '30's as a formalization of the informal notion of "potentially mechanizable" deduction, which was at the core of foundational analysis, since late '800, and resulted from the happy marriage of a Formalist and Intuitionistic understanding of effectiveness. For Peano, Hilbert and many others, foundation and certainty, in Mathematics, had to be found in stepwise deductions within "finitistic" system. For these leading mathematicians, the abstract and symbolic nature of Mathematics, but rigor as well, were identified and all coincided with *formalism* or with the "yet-to-be-specified" notion of *potentially mechanizable process*. At the core of it, there was the newly invented formal Theory of Numbers, Peano Arithmetic (PA), based on Peano-Dedekind formal rule for induction.

Along the lines of Leibniz's philosophical intuitions, Peano's "pasigraphy" is an early example of formal language, where finite strings of symbols were manipulated in a way feasible by an (abstract) machine. Peano, Padoa, Hilbert ... were actually referring to an mathematical machine, as nor the steam engine, nor the petrol engine, but not even the computing devices by Babbage and Power could be the reference for the notion they had in mind. These concrete machines were the paradigms of changing times, but these pioneers

were thinking to a mathematical device yet to be invented, an absolute tool for a deductive mind.

Such a tool was fully described in the '30's. Its history is rich of independent inventions. Curry anticipated the notion of computable function by a simple language of combinators [16], proposed for the finite manipulation of strings, on the grounds of an algebraic theory due to Shoenfinkel. Herbrand, in 1930, see [33], had a mathematical definition of a class of functions which were meant to represent logical deduction. Gödel ([32]) gave another characterization for the analysis of arithmetical deduction (i. e. for the metatheory of PA). And he encoded this metatheory within PA itself by two fundamental steps : he "Gödel-numbered" formulas by integer numbers and represented his class of (partial) recursive - or computable - functions into PA (a much harder task than programming in Assembler). By this operation Gödel ended Hilbert's dream of a final solution to all foundational problems : prove formally the consistency of PA, prove that all arithmetic statements can be decided and ... mathematically live in the fregean Paradise of ideal (infinitary) objects.

As a matter of fact, the (formalized) mathematics of ideal objects, Geometry and Analysis, had been encoded into Arithmetic by [36] and by Cantor and Dedekind, respectively. Thus, foundations, certainty and knowledge, for Hilbert, could be all transferred into the provable consistency and completeness (decidability) of Arithmetic. But, by encoding the metatheory of PA into the theory, Gödel showed that the encoded metatheoretic proposition of consistency is an undecidable statement of PA. If his class of computable functions were fully general (i.e. it convincingly represented the original, but informal idea of "potentially mechanisable - deductive - process"), then this one fact was enough to negate both Hilbert's conjectures. And it was so. A few years later, in [64], Turing gave the most convincing, abstract notion of formal machine ever possible. Key idea : distinguish between hardware and software. Here is the reading/writing head and the tape, there are the programs. Formal Logic finally had his locus for absolute certainty : move right, move left the head, read/write 1 or 0 on the tape ..., according to a program written in a finite number of well-formed lines, possibly encoded by 0's and 1's. This is the action of the leibnizian man in the act of deducing, the human computor, the definitive Calculemus : no further reduction is possible to more elementary steps. And yet the Machine is fully expressive : Curry, Herbrand, Gödel computable functions, but Church's ([14]) and Kleene's as well can all be programmed over Turing Machines (TM), [41]. At once, an "absolute" was proposed, in Gödel's words, and Computer Science was born. The absolute, logical Machine.

Invented for the purposes of formal Logic, TM's engendered the early generations of computers, the sequential machines that dominated the scene for very long. The issue of space, and of its geometric intelligibility, had been programmatically excluded from the foundational project, all centered on Logic since Frege (see [63], [47] and their references for more on this) : the result is that the expressive power of TM's, a formal machine for sequential deductions, does not depend on its being in space, e.g. it is independent from the Cartesian dimension of the device (TM's, in any finite dimension, compute the same class of functions). We will hint how this affects the geometric understanding of computations, briefly presented below (its so called "Denotational Semantics") as well as the novel challenges posed by today's modern computing, that of *concurrent* systems. These systems happen to be *distributed in space*, as networks of computers. As a matter of fact, classical (and intuitionistic) Logic and their formal machines are no longer sufficient to deal with the

new mathematical challenges posed by actual computing over concurrent systems. TM's are logical machines, not physical ones, as no relevant physical process is independent from the dimension of space (many depend only on it). And Mathematics, Geometry, proves it.

1.2 Dimensions

Computations take place along time, a linear time scanned by the clock of a Turing Machine (TM). This absolute, Newtonian understanding of time will be in the background of the approach to computations mostly followed here. But what about the dimensions of space ? There are two (equivalent) ways to understand its role in Computability Theory. First, some simple theorems prove that one linear tape, one head or finitely dimensional tapes and heads TM's have all the same expressive power, modulo some (low) complexity issue. That is, they compute the same class of functions, the Partial Recursive Functions (PR), as functions from ω to ω , the integer numbers. Second, once given PR, how to define computable functions over ω^n, for any n ? Easy : a many arguments function is **computable** iff it is so *componentwise*. Equivalently, in order to define computability from ω^n to ω^m, encode ω^n onto ω, in some effective and bijective way, and use PR. These two facts are a symptom of the irrelevance of physical dimension in classical computability.

As a consequence, all categories for computations below may be easily shown to be Cartesian, that is, to be closed under finite products. But more than this, functions of any finite number of parameters are all treated, computationally, in the same way. The componentwise nature of computability will be better understood by the Cartesian Closure of some of these categories, a stronger property (see below). These categories are an effective frame for dealing with functions of higher types (functions of functions ...) and on abstract structures. This is needed, in a mathematical treatment of computations, as programs may compute with (encoded) programs as input and, more generally, with abstract types of data. That is, a TM may take (the program of) another TM as input : mathematically this yields a functional acting on a function.

Of course, componentwise computability of PR (and the irrelevance of dimensions) presents a first mathematical difficulty, when dealing with them and their function space by analytic tools : continuity is a topological notion and global continuity (w. r. to several parameters) is not entailed by componentwise continuity. Moreover, Cartesian dimension is a topological invariant. That is, if one embeds discrete structures into topological continua as space manifolds (in Riemann's sense), then the simple isomorphisms above between arbitrary finite dimensional discrete spaces, between ω^n and ω, say, cannot be extended to continuous isomorphisms, even not to isomorphic embeddings. This is so, at least, for traditional approaches to the continuum, which refer to physical continua and yield a physically plausible (and classical) notion of neighborhood, i.e. where different points in space may be separated by the topology (or the topology is at least T_2, i.e. Hausdorff).

Thus, the invention of suitable, weakly separated topologies was required, which serve the "embedding purposes" and is fully motivated by the theorems in section 2.2. The Mathematics is very interesting and has had many relevant applications, in particular to the Theory of Programming, but the loss of the sense of space is not less relevant. Interestingly enough, this came out to the limelight from within Computer Science itself, as it will be hinted below.

This lecture is not meant to be a complete survey of "Geometry in Computing", but it mainly focuses on a (rather sucessful) topological approach. For this purpose, we will first survey the general frame for computability in some mathematical spaces and try to stress the need for topological tools for their sound treatment (sections 2, 3). Then, two specific applications of topological methods will be given. They are mathematical applications, and properly "analytic" ; that is, a purely combinatorial statement will be proved by an embedding of discrete structures into topological continua (sections 4, 5).

The limits of this approach will be hinted as well, in reference to the novel needs for Logic and Computing (see the Intermezzo). The point is that these topologies are directly (and provably) related to classical computability, that is to the logico-mathematical theory born in the '30's, while both Logic and Theoretical Computing need today to reconsider the mathematical understanding of space as an integral part of their contribution to knowledge and applications. Turing Computability should be gradually considered as classical (Laplace's) Rational Mechanics in Mathematical Physics : a major start, but entirely superseded by the Mathematics (the Geometry) of Dynamical Systems, Relativity and Quantum Mechanics. Yet, the main ideas, born in between 1930 and 2000, are worth recalling, as they may yield some fundamental mathematical concepts. In Rational Mechanics, say, laplacians or hamiltonians, as geodetics and energy conservation properties, passed over to XX century Mathematical Physics, as key conceptual and mathematical invariants (with all due changes, of course). As for now, the notion of sequential computation, Type, Polymorphism, sub-typing, some limit constructions and their topological and categorical understanding seem to remain basic tools in Programs' Theory and Design, in spite of the dramatic changes Computer Science is undergoing. Future History will better tell.

2 Scott Topologies and Computable Functionals

It is fair to say that Myhill-Shepherdson Theorem [52], in Recursion Theory, is one of the earliest results which provided a "topological", though implicit, characterization of a purely recursion theoretic notion of functional (or function with function argument). The topological content of this relevant theorem was made explicit and generalized by Dana Scott in the early '70's, for an analysis of typed and type-free computations (see below). Since then, the Scott topologies and their numerous variants have had a major role, both in higher-types Recursion Theory (where Ershov later rediscovered the same structures, within an original frame) and, more importantly, in the Denotational Semantics of Programming Languages.

Higher Type Recursion and Denotational Semantics have different but related motivations. First, the intuitionistic analysis of propositions naturally yields higher type computations : as it will be better explained below, alternating quantifiers are understood as (effective) functions ("for all x, there exists y ... " yields a computable function from x to y - this is the sense of the "constructive existence" of y), then further alternations of quantifiers yield functions of functions ... (see section 4). Second, Programming Languages need types, as they work on various "types of data", in particular of any finite order or level, as programs may take numbers as inputs but programs as well or even programs acting on

programs .... In either case, in Logic and in Programming, a crucial mathematical issue is posed : in Mathematics, when working with functions, as sets of input/output values, one has to deal with infinite objects. And the royal way out towards infinity, in constructive systems, is provided by approximation. But approximation means *order*, and *topology*.

2.1 Continuous Domains

Recall that computations, since Turing's early results, are essentially partial (the halting problem is undecidable or there exists partial computable functions that cannot be extended to total ones). There is now a natural way to order partial functions as input/output devices : they can be partially ordered by graph inclusion (that is, "smaller" means "less defined"). This is the underlying intuition that motivates the constructions below.

Consider a partial order $D = (D, \leq)$. For the purposes of an analysis of approximation, we need to generalize the notion of "chain" and "limits" in total orders to partial ones. A subset A of D is **directed** if for x and y in A, there exists z in A such that $x, y \leq z$ (a chain is particular directed set, of course). The limits will be provided by the suprema of directed sets, $\sup A$, whenever they exist.

Definition 2.1 *Let x and y be elements of D, then x is* **way below** *y (notation : $x \ll y$) if : for any directed set A in D, if $y \leq \sup A$, then $\exists a \in A\ x \leq a$.*

Clearly, the relation $\ll$ is finer than $<$, and one has : $x \ll y$ and $x' \leq x$ and $y \leq y'$ imply $x' \ll y'$. Notation : $\Uparrow a = \{b \in D \mid a \ll b\}$; $\Downarrow a = \{b \in D \mid b \ll a\}$.

Definition 2.2 *D is a* **Continuous Domain** *if, for each $a \in D$, the set $\Downarrow a$ is directed and $a = \sup \Downarrow a$.*

Consider a continuous domain D. Then $\emptyset \cup \{\Uparrow a \mid a \in D\}$ trivially forms a basis for a topology (the order topology, i.e. $x \leq y$ iff for all open U, $x \in U \Rightarrow y \in U$). Call this topology, the **Scott topology**. A subset D_0 is a **basis**, if, for each $a \in D$, the set $\Downarrow a \cap D_0$ is directed and $a = \sup \Downarrow a \cap D_0$. Clearly, also $\emptyset \cup \{a \mid a \in D_0\}$ is a basis for the Scott topology. If D has a countable base, then D is **ω-continuous**.

Note that the notion of approximation (or *neighborhood*) is given "from below" and that Scott open sets are upward closed : moving up, along the partial order, gives more and more defined (better approximated) elements ([60], [61]).

Example 2.3 *Let $P\omega$ be the powerset of ω , the set of natural numbers, and P the set of partial (number theoretic) functions. $P\omega$ and P are ω-continuous domains by set inclusion, where $a \ll b$ for some b iff a is finite iff $a \ll a$ (finite, in P, means "function with a finite graph", or equivalently "with a finite domain" - of definition ; approximation is graph inclusion - thus, the total functions are the maximal elements).*

In general, call **algebraic** any element a in a continuous domain, that satisfies the $a \ll a$ relation. This is a key general notion of "finiteness" that will be extended to functions and functionals in higher types : it yields the useful properties of "finite elements" in more general

settings, where there are no (properly) finite elements. As a preliminary exercise, consider two continuous domains A, B and the collection $CD[A, B]$ of continuous functions (w.r.t. the Scott topology) ; give to $CD[A, B]$ the structure of a continuous domain by taking the pointwise convergence topology on it. (When $A = B = P$, finite elements in $CD[P, P]$ are non-obvious but useful functionals - for example, they can be effectively enumerated, see below).

However, many interesting domains have no algebraic elements. Here is one that will be used later on, for further applications of Domain theory.

Example 2.4 *(The* **upperspace** *of* $[0, 1]$*). Consider the collection* $I([0, 1])$ *of the compact intervals in the real interval* $[0, 1]$*. Let* $D = (I([0, 1]), \supseteq)$*, that is* $\leq$ *is the* reverse *inclusion. Then :*

- $(I([0, 1]), \supseteq)$ *is an* ω*-continuous domain (base : the intervals with rational end points).*
- $[a, b] \ll [c, d]$ iff $[c, d] \supseteq]a, b[$
- *the singleton intervals are the largest elements and* $[0, 1]$ *is the least one.*

The intuition should be clear : in $I([0, 1])$*, the real points are approximated from below by decreasing the size of the intervals containing it. No element of this domain is algebraic. Note also that a base for the Scott topology is given by* $\{[a, b] \mid [a, b] \subseteq O\}$*, when* O *ranges over the open sets (of the real topology over* $[0, 1]$*). Clearly, the induced topology from* $I[0, 1]$ *over* $[0, 1]$*, the set of maximal elements, is the real topology. Thus, approximation from below gives the reals as fully specified elements (the total ones, in a sense), with their familiar topology.*

2.2 Myhill-Shepherdson Theorem

Consider now an ω-continuous domain D. Then its basis D_0 can be enumerated by a bijection $e : \omega \to D_0$, say. The domain is called **effective** if $e(n) \leq e(m)$ is a decidable predicate in n and m. This is a key property as for effectiveness, as it clearly applies to our leading examples above, in particular $P\omega$ and P, and it allows to define the collection of computable elements in an effective domain : just take the limit (the sup) of all effectively enumerable directed set in D_0. That is, consider the directed subsets A which possess a recursively enumerable (r.e.) set of indices, i.e. $\{n \mid e(n) \in A\}$ is r.e.. Their suprema, if they exist, will be the generalized computable functions and functionals in (higher) types or in abstract structures. In $P\omega$ and P, these sets give exactly the r.e. sets ($RE \subseteq P\omega$) and the partial recursive functions ($PR \subseteq P$). In $(I([0, 1]), \supseteq)$ they are the intervals with computable reals as end points (a real is computable, when it is "effectively generated", as a sequence). Thus, if one starts by considering TM's and their computable functions (PR or RE), we are now seeing them as embedded into suitable continua (P or $P\omega$), endowed with an order topology (note that both $(P, \subseteq)$ and $(P\omega, \subseteq)$ contain chains with the order structure of the reals).

Is this construction of a category of domains, really "natural" ? At least a theorem with two important corollaries proves it, by establishing a nice correspondence between topologies and computations. In order to state them, observe first that it is easy to enumerate the limits or sups of r.e. directed sets of the base, if they exist : just use the indices of the r.e. set, which enumerate the directed set below it. Each of these limits, of course, will have infinitely many indices, at least as many as the indices of enumerating r.e. set. In an effective

ω-continuous domain D, call D_c the set of the computable elements in D, and $w : \omega \to D_c$ this (canonical) enumeration of D_c. If all limits of the r.e. directed sets of the base exist in D, call then **effectively complete** such an effective domain. Then one has :

Theorem 2.5 *In an effectively complete ω-continuous domain D, consider a canonical enumeration $w : \omega \to D_c$ and $B \subseteq D_c$. Then, if $\{n \mid w(n) \in B\}$ is recursively enumerable, one has that B is an open set (in the Scott topology). Conversely, for any basic open set $B \subseteq D_c$, the set $\{n \mid w(n) \in B\}$ is recursively enumerable.*

The first implication is a non-obvious fact, with several consequences (the converse is easy). Observe first that there is no apparent reason by which a recursion theoretic notion, the recursive enumerability of a set of indices, should lead to topological openness. Look at it in $P\omega$ or P, where it was first proved by Rice and Shapiro (see Theorem 14.XIV in [57]) by a non-trivial argument : its key lemma, later called of "effective discontinuity", has a relevant logical-intuitionistic meaning, and it may be also reconstructed from the proof of Myhill-Sheperdson Theorem, see corollary 2.7 below (the generalisation to domains given here is easy). The theorem above says that if you take a collection of r.e. sets or of partial recursive functions, then their semi-decidability gives a certain order or topological structure to the set of sets or of functions, considered in extenso (as set-theoretic relations, e.g. as elements of $P\omega$ or P). An immediate consequence is that a property of r.e. sets or partial recursive functions is decidable iff it is trivial. More generally :

Corollary 2.6 *In an effectively complete ω-continuous domain D, with $w : \omega \to D_c$ and $B \subseteq D_c$, $\{n \mid w(n) \in B\}$ is recursive iff B is empty or coincides with D.*

Proof. Observe that an effectively complete ω-continuous domain has necessarily a least element, the sup of the empty set (which may be effectively enumerated by the empty set of numbers). Assume then that $\{n \mid w(n) \in B\}$ is recursive, i.e. that it is r.e. with its complement, then the set and its complement are both Scott open, by the theorem. One of them must contain the least element and thus, coincide with D. The converse is obvious. □

The relevance of the Corollary should be clear, even in the basic case of $P\omega$ or P, where the result was originally given in [58]. In summary, no non-trivial property of programs is decidable (Corollary 2.6), and the semi-decidable properties are "very few" or "topologically very structured" (Theorem 2.5). In particular, we cannot decide nor semi-decide whether a program is correct, that is if it computes a given function or a function in the desired set. Large part of Theoretical Computer Science has been dedicated, in the last 30 years, to the analysis of program properties and to the invention of methods for partial correctness, exactly because these properties are not decidable (and very rarely partially decidable). Type Theory provides some of these methods and we will discuss it as well as the topological structures that it inspired.

The approach by Scott domains easily gives a notion of computable functional or function in all finite (higher) type. Call now SD the category of (possibly, effectively complete ω-continuous) Scott domains and continuous maps as morphisms. Observe then that any finite Cartesian product of these domains is still a domain (in each of the intended categories : continuous, ω-continuous, effectively complete ω-continuous). That is, they are all Cartesian

Categories. Consider further two (effectively complete ω-continuous) domains D and D', then one may show that the set of continuous functions from D to D' yields an effectively complete ω-continuous domain, $SD[D, D']$. The construction of $SD[D, D']$ is (relatively) easy : partially order this set of functions pointwise, observe that the Scott topology is the pointwise convergence topology and give all the required enumerations.

But something more happens : products and function spaces nicely commute. That is, there is a natural isomorphism, in the sense of Category Theory, between the functors $SD[D\times -, C]$ and $SD[D, C^-]$ (where C^- is the contravariant exponent functor). In technical terms, these categories are Cartesian Closed. The natural isomorphisms guaranty that the higher type sets, e.g. $SD[D', C]$, are internally represented by the exponents; that is, $C^{D'}$ is an object of the category that "corresponds" to the set of function $SD[D', C]$. Thus, in each higher type one has, *internally*, continuous (and computable) elements, continuous (and computable) functions on them and so on so forth. Moreover, a function of several variables may be seen as a continuous function of one variable at the time (or, componentwise continuity implies global continuity, as $SD[D, C^{D'}]$ and $SD[D \times D', C]$ are isomorphic).

However, does this approach yield a sound notion of higher type computation ? In order to prove this, consider a more elementary or purely computational approach against which we can check the one above; an approach with no reference to topologies and approximations. Given two enumerated sets A and A', i.e. two countable sets and surjective maps (enumerations) $w : \omega \to A$ and $w' : \omega \to A'$, there is an obvious notion of"effective transformation" $F : A \to A'$. Just define F to be an **effective transformation**, if there exists a total recursive function f such that $F(w(n)) = w'(f(n))$. That is, F may be effectively computed over the indices. This category EN of enumerated sets and effective transformations is Cartesian but, and this is the rub, it is not Cartesian Closed. In short, there is no natural way to turn into an enumerated set the function space of two enumerated sets, or, in general, this construction does not commute with products, in the sense given above by a natural isomorphism.

The solution follows from the construction of Scott Domains. Consider the category CSD whose objects are the computable parts of effectively complete ω-continuous domains. That is, the computable elements D_c , D'_c in Scott domains D and D', enumerated by $w : \omega \to D_c$ and $w' : \omega \to D'_c$, respectively. By the construction above, this is Cartesian Closed, if one takes as morphisms the continuous and computable functions (as elements of the higher types). Then one has :

Corollary 2.7 *CSD is a full sub Cartesian Closed Category of EN.*

Proof. The only delicate point is in the proof that the embedding is full. But, by Theorem 2.5, the effective transformations on enumerated sets are continuous (the inverse image of an r.e. set by a total recursive function is r.e.). □

This fact shows that a purely recursion theoretic, or combinatorial, definition of higher type function (by using computable functions over the "indices", i.e. within EN) can be extended to all finite higher types. The definition actually commutes with products and yields the required uniformity in one or several parameters (componentwise computability and continuity), categorically called Cartesian Closure. This is obtained by embedding partial (recursive) functions into suitable topological spaces and by using continuity of functions

and pointwise convergence topologies on their function spaces. Corollary 2.7 is a (broadly) generalized version of Myhill-Shepherdson theorem.

It should be clear that we are not following history here. As a matter of fact, Rice theorem was proved first, as a consequence of the halting problem ([58]), then Myhill-Shepherdson and Rice-Shapiro theorems followed, but just over $P\omega$ and P, with no explicit reference to topology (see [57]). Note finally that a function on sets or functions (a type-two functional) transforms or "reduces" a set or a function to another : Myhill-Shepherdson functionals, say, define "enumeration reducibilty", while further "reducibility notions" are given in purely recursion-theoretic terms (Turing reducibility, or reducibility by oracles, truth-table reducibility ..., see [57]). Yet, as proved in [7]), by (small) changes in the (induced) topology on $P\omega$, one may characterize these notions of "recursion-theoretic reducibility". Once more, a "geometric" perspective unifies notions and results. And, more importantly, Scott Domains gave the general mathematical frame and extended the results for type-two recursive functionals to the so-called Denotational Semantics of typed and type-free Functional Programming Languages, see section 3.

2.3 Total Functionals, Filter Spaces and Real Numbers

As already mentioned, the notion of "partial map" is fundamental in Recursion Theory, in view of the early results on undecidability. Attempts to work directly with total maps lead to complex syntactic approaches. The reason is first due to the impossibility to enumerate effectively the total recursive functions; thus only some subclasses (e.g. the primitive recursive functions) bear some interesting recursion theoretic analysis. As for higher types, Geometry, by an analysis of convergence, spells out the reasons for the mathematical difficulties. In [37] an elegant approach is proposed by Filter or Limit Spaces, à la Kuratowski. A Filter Space is a set with a local "notion of convergence" (for each point, a filter of filters of subsets); the category Fil has Filter Spaces, as objects, and, as morphisms, all maps transforming converging filters (towards a point) into converging filters (towards its image). Any Topological Space is a Filter Space with, at each point, the filter of neighborhoods as least filter (of convergence); this yields an adjunction from Top into Fil. Yet, not any collection (filter) of converging filters has a least element; that is, there are interesting objects of Fil, which are not in Top. For example, the type two *total* recursive functionals, formally defined by Kleene by 9 extremely complex axioms, [42]. By Hyland's "limit approach", some feasible (and elegant) Mathematics begun also on hereditarely total maps (see [55]). More recent consequences use this very weak notion of convergence and the following approach to computability over metric spaces.

In example 2.4 above, the reals, with the Euclidean metric, were embedded, as maximal elements, into a Scott Domain (the so-called "upperspace"). This is a paradigmatic treatment of classic structures within a suitable frame for computability and opened the way to some relevant effective analysis of continua. Computable reals, say, are exactly the maximal computable elements of the upperspace, as an effectively complete ω-continuous domain. The work by Edalat and others (see [19], [20], [24]) allowed to embed several structures for Effective Analysis, including some Geometry of Dynamical Systems, into Scott's categories, where closure properties under higher type constructions and existence of effective limits have been used to revisit several known results, in a constructive frame (effective versions

and/or improvement of Banach fixed point theorem, Hutchinson's theorem, existence of Markov operators, see [23] for a survey). A general idea, for example, is to look at spaces of open balls in a metric space as a Continuous Domain, partially ordered under reverse inclusion, as in the upperspace over $[0, 1]$ (example 2.4) ; then Ruelle's notions of "attractivity" and "invariance" can be described as convergence and fixed points and dealt with in a constructive way, also in higher types (see below and [22]).

A blend of the two approaches above has been recently proposed in [17]. The filter structure on reals and, hereditarely, on their functions and functionals, gives a sound notion of effectiveness in higher types (functionals and measures over the reals) by the combined use of weak forms of convergence (filters) and strong algebraic-geometric properties (convexity and the vectorial structure of the reals and their continuous functions). But this work belongs already to the new century.

3 Types

As already mentioned, Computability Theory was born by the encounter of formalist and intuitionistic views on foundations. In modern terms, this may be understood by the "Types as Propositions" analogy (also known as the Curry-Howard isomorphism below). In [14] a simple formal language for computing was defined, the type-free λ-calculus. The purpose, as for Herbrand, Gödel, Turing and the others in the '30's, was the understanding of formal deductions as effective/finitistic procedures. Church's elegant theory of computable functions only relies on two notions : given the formal definition of a function by a term, bind (abstract) a variable by a λ ; apply a term to another term. Briefly, $\lambda x.f(x)$ is an (explicit) function of x and when applied to a, say, it gives $f(a)$.

More formally : variables are terms ; if M is a term, then $\lambda x.M$ is a term ; if M and N are terms, then MN is a term. Warning : also xx is a term ! This immediately produces a fixed-point operator : $Y = \lambda y.(\lambda x.y(xx))(\lambda x.y(xx))$. As a consequence, Curry gave a simple antinomy within the calculus, which originally included an negation operation, Neg. As an informal paraphrasis of it, note that $YM = M(YM)$ for each term M, thus $Y(Neg) = Neg(Y(Neg))$ contradicts any reasonable property for negation or complement. It was simple to give a solution to the antinomy (yet another obtained by a diagonal construction) : avoid complementation in the calculus, as this remains very expressive. As a matter of fact, by a simple encoding of numbers as terms, type-free λ-calculus computes all (Turing) computable functions : it actually served as a tool for proving the equivalence results between the various formalisms, in '35-'36 (the Y operator above yields recursion). The lack of negation may be then understood as the fact that recursively enumerable sets are not closed under complement, an easy consequence of Turing's undecidability of the Halting Problem.

Yet, there is another relevant way out from antinomies : give types to variables, and thus to terms as suggested by [15]. Actually, one can introduce and eliminate types and define terms at once, by two rules :

$$\frac{\begin{array}{c}[x:A]\\ \vdots\\ b:B\end{array}}{(\lambda x:A.b):A\to B}(\to I) \qquad \frac{a:A \;\; c:A\to B}{ca:B}(\to E)$$

The reader will recognize in $(\to E)$ the classic 'modus ponens' : if A and A imply B, then B. While the rule $(\to I)$ transfers into the language of formulas the metalinguistic deduction : that is "from A infer B" is transformed into an implication, $A \to B$. By this very construction, the types of terms are also propositions of a simple implicative calculus. Just add products "$\times$" as conjunction and you have Intuitionistic Propositional Calculus, with the obvious rules that characterize the Cartesian projections. This was Curry-Howard's remark in the '60's that gave full logical meaning to Church's typed version of λ-calculus. The idea is that terms living in types are *proofs* of these types, seen as propositions : $(\lambda x : A.b) : A \to B$ is a proof of $A \to B$ in the precise sense that it takes a proof of A and transforms it into a proof of B, as expressed by $(\to E)$.

Note now that, by definition, several argument functions in λ-calculus are given componentwise. If M (possibly) contains x and y, the function explicitly depending on them as variables is obtained by successive binding by λ : $\lambda x.\lambda y.M$. Thus, as for TM's or any other calculus for recursive functions, computability is defined componentwise ; actually, in this calculus, the types $A \times B \to C$ and $A \to (B \to C)$ are provably isomorphic (there exists invertible terms in either direction). Clearly, this isomorphism has a precise logical meaning : as propositions, it is easy to prove that $A \times B \to C$ and $A \to (B \to C)$ are equivalent. This is a basic example of the connection between computations (as terms) and proofs, thus between types and propositions ; their isomorphism is constructively given by the terms proving their equivalence as propositions.

And here come again the Cartesian Closed Categories (CCC's), as this isomorphism is the natural isomorphism, which defines them (see above). Thus, Logic and computations force this topologically un-natural feature : global continuity is equivalent to componentwise continuity, a mathematical translation (or semantics) of the computational and logical isomorphism above. Yet, the construction of Scott Domains gave a simple answer to the problem, in topological terms. And the construction was far from arbitrary : functions, as infinite objects, are naturally partially ordered by graph inclusion (and this yields the pointwise convergence topology). Finite functions (functions of finite graphs) provide the finitary (algebraic) approximating elements, hereditarily extended to higher types by the categorical construction ; theorem 2.5 and its corollaries above further justify it. Clearly the price to be paid is that "topological approximation" is defined only from below : these topologies are order topologies, and not Hausdorff.

Besides this general motivation for continuity (approximation), there is a further motivation for Scott topologies, at least up to these days. The type-free λ-calculus is expressive enough to compute all partial recursive functions : this is easily show by implementing recursion within the calculus, by Curry's fixed-point operator $Y = \lambda y.(\lambda x.y(xx))(\lambda x.y(xx))$. Recall that, in this calculus, functions may be inputs as well (any term can be applied to any other term : there no difference of type). Formally, the domain of terms, D, is isomorphic to the domain of functions on terms, $D \to D$. But how to find a solution D to the equation

(isomorphism) $X = X \to X$?

By Cantor's cardinality theorem, this is far from obvious, if we want D to be different from the singleton set (e.g. to contain all natural numbers and computable functions). Well, in Scott Domains, by an inverse limit construction, one can find a solution to this equation, as isomorphism between a domain and its function space (the continuous functions or morphisms in SD), called the D_∞ model in [61] (see [6], [1] for surveys and references). A simpler, but weaker, solution, on $P\omega$ actually, is given in [Scott76] (for general category-theoretic and set-theoretic analyses of continuous domains and their equations, see [62] and [45], respectively). It is worth noticing that, so far, while there are plenty of interesting Cartesian Closed Categories (thus, models of the *typed* λ-calculus), no model has been given of the type-free calculus, without some use of continuity and topological notions.

Clearly, the very simple order topology of these models has an immediate advantage. Recursive functions are defined by ... recursion ; that is, a function f is usually given by a formal expression, a program, a "context of definition" F and an equation $f = F(f)$. Thus, partially ordered spaces, which are (effectively) complete, easily yield a solution f_0 of the equation : once interpreted F by a continuous, thus monotone map F, take the least upper bound of the chain $\perp = F(\perp) = F(F(\perp)) = \dots$, where $\perp$ is the least element of the partial order. The (continuous and computable) functional Fix, defined by $Fix(F) = \sup_n F^n(\perp)$, provides a simple mathematical interpretation of the formal operator Y above. Thus, a solution of the formal equation $f = F(f)$ is given in these topological spaces as $f_0 = Fix(F)$

As for the type equation $X = X \to X$, or more generally $X = \Sigma(X)$, where Σ is a type constructor or a recursive definition of data types, one may try to use a similar techniques, at the level of the category, by interpreting Σ as an (endo-)functor. However, the key case above - and most interesting ones - use functors that may be covariant in one argument (the second one), but contravariant in the other : thus, no obvious extension of the monotonicity argument can be given. Scott's '72 paper, by an original inverse limit construction, started the Denotational Semantics of Programming Languages, as the mathematical investigation of recursive definitions of functions and, more importantly, of data types : this crucial practice of programming found by this a proper mathematical frame, while the first (topological) model of type-free λ-calculus was given, i.e. a solution of the equation $X = X \to X$.

In conclusion, types of (typed) λ-terms are propositions ; their mathematical meaning is that of objects in categories with some closure properties, that relate products and function spaces. Some of them contain non-trivial solutions of the equation $X = X \to X$. These correspondences (types as propositions as objects of CCC's) greatly enriched Proof Theory as well, in particular since the work of Girard and Martin-Löf, in the '70's. On the side of applications, type-free and typed λ-calculi have been the paradigms for Functional Languages, one of the main programming styles, for more than 30 years. The main dialects of LISP, born in the '60's, directly derived from the type-free calculus ; Edinburgh ML, an extension of the typed calculus of the mid '70's, started a major flow of programming languages. By the notion of *polymorphism*, due in Proof Theory to Girard's second order λ-calculus, novel and rigorous reusability tools moved into programming (see [30], [13] for surveys). Moreover, their mathematical semantics provided connections to further non-trivial constructions, which also originated in Geometry : Toposes.

In short, polymorphism may be understood as the possibility, for a function (a term) to live in many types. The identity function, for example, belongs to $X \to X$, for any X ; or,

one should be able to write just one program computing it, in all types. The idea is that types may include "variable types". Then, both types and terms may depend on types, as variables : e.g., $(\lambda X.\lambda x : X.x) : (\forall X.X \to X)$ is the polymorphic identity function, or the program for this function may be "reused" over all types. More generally, if A is a type, also $\forall X.A$ is a type, for X ranging over types (formally : $(\forall X : Type.A) : Type$). Thus one may define a type by a quantification over the very collection of types that is being defined, see [30]. This challenging impredicative feature is the reason for the expressivity of the calculus (Girard's system computes all recursive functions which are provably total in II order Arithmetic, and thus it proves the consistency of - II order - Arithmetic), as well of its success, due to the idea of "reusing" as polymorphisms or variable types, with no restrictions (no stratified universes, for example).

Girard's second order calculus (with type variables) may be interpreted over toposes (where variation is described as "fibration"), with a non-obvious closure property, which may be informally expressed as follows. The intended category must be closed under fibred or indexed products *over itself*. That is, it must contain, as an object, the result of a product indexed over the entire collection of objects ("small completeness"). The first construction of this kind was proposed by Moggi, in 1985, and developed by many since then ([50], [39], [56] ; see [4] for an approach via "internal categories" and more references). Moggi's idea may be accommodated into several realisability toposes, a natural generalization of computability in categorical/ intuitionistic frameworks, which contains various versions of Scott Domains. Their logic is essentially intuitionistic and they allow strong closure properties exactly because they yield "very few morphisms" (only the "constructive ones"), see [10] for surveys and recent advances.

Realizabilty toposes do not necessarily relate to continuity notions. However, their origin is due to Grothendiek's idea of topos, which extended the gluing technique of Riemann, for differentiable functions on manifolds, to more general structures, see [51]. Lawvere made the connection to Logic, [44]. The main theorem in [50] uses Lawvere's understanding of (second order) universal quantification as right adjoint to the diagonal functor (note that existential quantification is understood as the left adjoint : a very elegant "geometric" understanding of logical quantification, by a symmetry).

Intermezzo : Logic, Concurrent Processes and Geometry

Logic and Sequentiality

As mentioned in the Introduction, the logico-formal origin of Computability Theory is reflected into the topological structures that were first proposed to understand it (as already hinted, since [26], foundations only relied on "logic judgements", with the exclusion of "geometric judgements", as hypothesis grounded on key regularities of space, in the sense of [59], [65] ; see [47] for a discussion). The componentwise analysis of continuity, as Cartesian Closure of the intended categories, the weakness of the topology (T_0 separation) ... are all symptoms of this "lack of attention" to physical space (and its cartesian dimension) which is typical of this theory. The sequentiality of (classical and intuitionistic) logical deduction is at the core of it. TM's are conceived as a human computor sequentially writing/erasing/moving 0's and 1's. And theorems prove that λ-calculus is a sequential system (see [9], [3]). Also the

results mentioned in subsection 2.2 have an essentially sequential nature. A close analysis of the proof of the "effective discontinuity lemma" (see the comments after theorem 2.5) shows that its proof relies on the sequential generation of an r.e. set : a bifurcation towards a superset or a finite subset is decided, with no other alternatives, along its effective production ; and this seems crucial to the construction. As a further remark, recall that the type-free λ-calculus may be considered as the cleanest functional theory of computations, as it represents all partial recursive functions with the simplest syntactic structure. Observe then that all its models force the strong isomorphic embedding that we considered as a paradigm of "being unrelated to physical space" (as dimension is a topological invariant) : in any Cartesian Closed Category, a solution D of the equation $X = X \to X$ is such that $D \times D$ can be isomorphically embedded into D (a simple consequence of the syntax or of the categorical characterization in [49]).

Yet, the deep links between Computability, Cartesian Closure and weak topologies, as concrete models for types and computations, have had lots of remarkable fall-out. Functional programming, since the design of (dialects of) LISP (such as Scheme, and their models) and even more as for Edinburgh ML, till the more recent versions of polymorphic programming languages, have been invented and enriched on the grounds of this mathematical understanding. As already hinted, some of the related ideas passed into Object Oriented progamming : for example, subtyping helped to formally describe inheritance (the features of programs may be inherited and reused, a further form of polymorphism, see [13]). But even Logic has been influenced by these structures. On the grounds of early work by Berry, in [28] it was remarked that one can give models for typed and type-free computations by considering a restricted class of continuous functions as morphisms, the stable ones. In the categories of Berry-Girard Coherent Domains, the key property of "finiteness", interpreted by the algebraic elements of Scott Domains, is nicely preserved, as proper finiteness also in higher types (there are only finitely many elements below a finite - algebraic - one ; see [3] for an account). And this was the beginning of new ideas in Proof Theory : Girard's Linear Logic was conceived by observing that stable maps include some interesting *linear* ones, [29]. In this case, then, the mathematical structure preceded and suggested the logic formalization : the linear maps interpret a novel, very fruitful, notion of implication.

Girard's work went further on, in this revision of classical and intuitionistic systems. Linear Logic originated an original analysis of proofs, in terms of their geometric structures : proof-nets describe rules and proofs also on the grounds of *symmetries* and *connectivity*. The logic machinery of the founding fathers is thus enriched by some fundamental geometric judgements that bring space back into Foundations, through the window of deductions. A further insight on the relevance of spatial organization in Logic is given in [31].

Concurrency and Geometry

It is clear that Turing machines may run in parallel, and that one machine can simulate any finite number of them. This is parallelism and its use largely enhanced complexity, but did not require any change in paradigm. The situation, in actual computing, changed dramatically since distributed systems and networks of computers came to the limelight. In these systems, possibly asynchronous processors concur in performing a task. That is, they exchange messages that may modify each individual computation, in a largely unpredictable

way, in particular in view of the lack of an absolute, Newtonian clock. Distribution in space is crucial, as well as a richer understanding of physical time. Their mathematics is changing accordingly. A wide effort is carried on in order to provide a suitable mathematical frame for Concurrency (see [2] for recent approaches and surveys). Categories play a major role, see [66], [43] among others, but we are far from a unified understanding of a very complex phenomenon, where computations are no longer compositional (the result of truly concurrent processes cannot be analyzed by composing the independent analyses of individual processes ; this radically departs from Turing's approach[1]).

Even access to stored data is becoming a major mathematical challenge : competing independent processors, in a distributed systems, may modify a database while accessing to it and, by this, influence the other computations. A recent geometric approach to this relevant and difficult problem uses non-trivial tools from Homotopy Theory. The idea is that the allowed computations may be seen as homotopic paths in a continuous phases space (see [34]). Fault tolerance, typically, is analyzed in terms of small continuous deformations along oriented paths. Space is finally back and true geometry follows. The point is that Computer Science is moving from the discipline engendered from pure Logic, an heritage of the spaceless foundations à la Frege and Hilbert, focusing only on sequential languages for deductions, Integer Numbers and their Arithmetic, to an analysis of actual computations by concurrent agents, distributed in space and computing along a dynamical understanding of time (time as scanned by irreversible and asynchronous bifurcations). The founding fathers of Logic, one century ago, opened the way, by their mathematical inventions, to the early logico-formal machines ; but these machines are changing nowdays. And it is time to move towards new ideas for the "next machine", possibly grounded on physical space and its mathematics. A good source of inspiration may be another relevant machine, rather unrelated to the digital ones, the animals' neural system (which includes ours, of course). This machine originated from *action in space*, along the evolution of life, and its elaboration of information seems more to be a dynamics of forms, of a geometric nature (from the tridimensional folding of proteins, as neuroreceptors, to the structure of neural nets - see the project "Géometrie et Cognition" in Longo's web page ; but this is yet another possible program for the new century).

4 The Hereditary Partial Effective Functionals

In this section and in the following one, two "finitary" statements are given, which have been proved by an explicit use of the continuous structures we previously introduced. It is fair to say that "analytic" methods have been used, for elements of domains are described by finite approximations, similarly to approximations by finite expansions of Fourier series in complex function theory. Of course, we do not yet deal with the width and the difficulties of number theory : the Theory of Numbers is the oldest and most basic topic in Mathematics,

[1] The number-theoretic functions computed by these systems are the usual ones, of course. The point is that, in Concurrency, what really matters is the data flow, not the input-output relation. Church Thesis, the thesis for the old absolutes of the 30's, for Computing (and Mind (!!)), may be soundly considered the "tar-pit" of modern (mathematical) computing, see [2].

since Pythagora's school. Computability Theory is surely not as old as that[2].

More precisely, this section presents an elementary (arithmetic) notion of higher type computation. The basic properties of the type structure are proved by an essential use of continuous domains for denotational semantics and generalized computability. In section 5 below, a (basic) fact for the solution of recursive equations between programs is discussed. This is done for our core functional language, λ-calculus. The proof uses a non-axiomatizable ("infinitary") consistent extension of λ-calculus and a continuous topological model. Unfortunately, some technical definitions are unavoidable.

Gödel and Kreisel's motivation for classical higher type recursion theory refers to the analysis of the meaning in constructive mathematics of Peano or Heyting arithmetic (see Kreisel's paper in [35]). Recall first that $\forall x \exists y\ A(x,y)$ (constructively) yields a function f from x to y, i.e. it corresponds to $\exists f \forall x\ A(x, f(y))$. More generally, following Kreisel's "no counterexample" interpretation,

$$\begin{array}{c} \text{from } \ \exists x\, \forall y\, \exists z\, \forall w\, A(x,y,z,w) \\ \text{write } \ \exists F\, \exists G\, \forall f\, \forall g\ A(F(f,g), f(F(f,g)),\, G(f,g), g(F(f,g).G(f,g))), \end{array}$$

where F, G, f, g are function variables in the due type (type 1, the type of number theoretic functions, as for f and g; (pure) type 2, as for the functionals F and G, since they are functions of functions). Then x and y are effectively recovered from f and g, if F and G are (higher type) total *computable* functionals, and so on so forth for further alternations of quantifiers. Thus the need to define this sort of computable maps, in any finite higher type.

Recall now that the partial recursive functions (PR) can be *effectively* numbered (notation : $PR = \{\phi_i\}_{i\in\omega}$); thus, also the universal function $u(i,x) = \phi_i(x)$ is in PR. By using this idea, in [48] a direct elementary approach was proposed to higher type partial functionals (the HPEF below).

The difference w.r. to Scott-Ershov approach (or Hyland's, for total functionals) is that the later first introduces the notion of continuity, possibly in abstract categorical settings, then defines the continuous and computable objects in any type, by using also PR. The HPEF instead are presented by an inductive elementary construction, only based on PR. Interestingly enough, the key structural property of the HPEF, which may be expressed in purely arithmetical terms, could only be proved by using continuous structures. This may be considered a further motivation to the invention of the topological category of domains.

4.1 The problem

For sets A, B, write $f : A \to B$ if f is a total map from A to B. By setting $\omega^{\perp} = \omega \cup \{\perp\}$, we may write $f : \omega \to \omega^{\perp}$ for any partial number-theoretic function ($\{\perp\}$ means "divergence"). As for higher types, integer (or, respectively, pure) types are defined by $n+1 : n \to n$ (or, respectively, $n+1 : n \to 0$), for $n \in \omega$.

Let $<,>$ be any effective coding of pairs in ω; that is $<,>$ codes a pair of numbers by a single number. Set $C^{(0)} = \omega$ and $C^{(1)} = PR$. In order to define $C^{(n)}$, the key idea is to

[2]The two theorems in sections 4 and 5 have been proved by [48] and [8], respectively. In [46] they were already presented in a unified perspective.

extend to a map from $C^{(n-1)}$ to $C^{(n)}$ the notion of Gödel-numbering, which is welknown as a map from $C^{(0)}$ to $C^{(1)}$

We will define then $C^{(n)}$, for all $n > 1$, as the set of **hereditary partial effective functionals** (HPEF) of integer type n, by induction on n, by using a set $C^{(n.5)}$ of maps from $C^{(n-1)}$ to $C^{(n)}$ (functions of "intermediate" type). The maps in $C^{(n)}$ will go from $C^{(n-1)}$ to $C^{(n-1)}$. The coding $<,>$ is extended to higher types in one of the several possible ways. In other words, the extented $<,>$ codes $C^{(n)} \times C^{(n)}$ by $C^{(n)}$, for any n, as it will be shown below.
Notation :

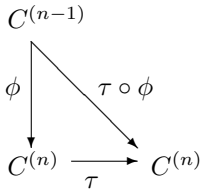
$$\lambda xy.g(x,y) \text{ is the map } < x, y > \vdash g(x,y).$$

Definition 4.1 *(i)* $C^{(n.5)} = \{\phi : C^{(n-1)} \to C^{(n)} \mid \lambda xy.\phi(x)(y) \in C^{(n)}\}$ *;*
ii) $C^{(n+1)} = \{\tau : C^{(n)} \to C^{(n)} \mid \forall \phi \in C^{(n.5)} \tau \circ \phi \in C^{(n.5)}\}$.

In order to see that this definition makes sense, let's check the types.

Assume that $C^{(n-1)}$ and $C^{(n)}$ are given. Consider first $\phi : C^{(n-1)} \to C^{(n)}$ and set $\psi(< x, y >) = \phi(x)(y)$, for $x, y \in C^{(n-1)}$ (check the typing for exercise : e.g., observe that $\phi(x) : C^{(n-1)} \to C^{(n-1)}$). Clearly $\psi : C^{(n-1)} \to C^{(n-1)}$. Then, by (i) in the definition above, $\phi \in C^{(n.5)}$ iff $\psi \in C^{(n)}$.

The following diagram should help in understanding how $C^{(n+1)}$, the "next" higher type, is defined in (ii), by using $C^{(n.5)}$, which has been defined in (i) from the given $C^{(n-1)}$ and $C^{(n)}$:



That is, $\tau \in C^{(n+1)}$ iff $\forall \phi \in C^{(n.5)}$ $\tau \circ \phi \in C^{(n.5)}$, as in (ii) in the definition.

Note that the pure type functionals $\{PC^{(n)}\}_{n\in\omega}$, i.e. for $n+1 : n \to 0$), may be also defined, by the technique above : just substitute $PC^{(n)}$ in the second instance of $C^{(n.5)}$ in the definition of $PC^{(n+1)}$ corresponding to definition 4.1 (ii) (see [48] ; 3.12)). In either cases, one only needs to know $C^{(0)} = \omega$ and $C^{(1)} = \text{PR}$, to start with. Apparently, though, it may seem that the full strength of Set Theory is needed in order to define the sets $C^{(n.5)}$ and $C^{(n+1)}$ above. As a matter of fact, a consequence of the results mentioned below is that, for all $n, C^{(n)}$ and $C^{(n.5)}$ are countable and can be effectively numbered. By this we do not need to look, say, at the set of *all* functions from $C^{(n-1)}$ to $C^{(n)}$ in order to define $C^{(n.5)}$, but a countable subset would do. A posteriori, this is all expressible in Peano Arithmetic.

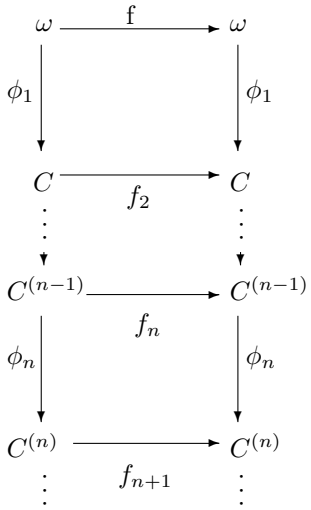
We want to know, now, what the $C^{(n)}$'s are and how they behave, for $n > 1$. It is easy to see that all acceptable Gödel-numberings of $C^{(1)} = \text{PR}$ live in $C^{(1.5)}$. It would be nice, say, if one could extend the notion of effective numbering to maps in $C^{(n.5)}$, for $n > 1$. In view of the s-m-n (iteration) theorem (see [57]), this corresponds to proving the property (P) below. This property says that each $C^{(n.5)}$ contains a function ϕ_n to which any other

function, in the same type, can be reduced (by a function $f_n \in C^{(n)}$). Function ϕ_n plays the role of "generalized Gödel-numbering".

$$(P) \qquad \forall n \geq 0 \ \exists \phi_n \in C^{(n.5)} \ \forall \phi \in C^{(n.5)} \exists f_n \in C^{(n)} \phi = \phi_n \circ f_n$$

Recall now that for all $f_{n+1} \in C^{(n+1)}$ one has $f_{n+1} \circ \phi_n \in C^{(n.5)}$. Then by (P), for some $f_n \in C^{(n)}$, $f_{n+1} \circ \phi_n = \phi_n \circ f_n$.

By this, HPEF may be visualized in the integer types by the following diagram :



In order (P) to hold, the ϕ_n's have to be surjective as well and, by this, preserve the cardinality of ω at higher types. Again, (P) is elementary and has a combinatory flavour.

4.2 The method

In this section we give the basic hints for the main result concerning the HPEF, with the only aim to display the use of continuous topological structures in the proofs. In [48], (P) has been proved by an heavy induction load. Note first that, by looking only at definition 4.1, the HPEF do not need to be ... well defined. As a matter of fact, $<,>: \omega^2 \leftrightarrow \omega$ is not trivially inherited at higher types. At first sight, one would just set $< f, g > (n) = < f(n), g(n) >$ as a definition for $<,>: C \times C \leftrightarrow C$ and so on. This is clearly wrong, for divergence of f or g cause essential problems. Set then $< f, g > (n) =$ **if** n is even **then** $f(\frac{n}{2})$ **else** $g(\frac{n-1}{2})$.

This is fine for $C^{(1)}$, but it doesn't automatically prove that
$< \tau, \sigma > (f) = < \tau(f), \sigma(f) >$
defines an (effective) coding of $C^{(n)} \times C^{(n)}$ into $C^{(n)}$, for $n \geq 2$. The proof of this goes inductively together with the proof of (P) and of two more facts, which need a few definitions.

Recall that a numbered set is a pair (D, e) where $e : \omega \to D$ is onto. $f : D \to D'$ is a **morphism** (of the numbered sets $(D, e), (D', e')$) if $\exists f' \in R \; f \circ e = e' \circ f'$ where R are the recursive functions. Define then $d : \omega \to D$ is an **acceptable numbering** of D if $\exists f, g \in R \; , \; e = d \circ g$ and $d = e \circ f$ (or, d and e are recursively equivalent). Now, one can enumerate D' from D : that is, $\phi : D \to D'$ is a **relative numbering** (of (D', e') w.r.t. (D, e)) if $\phi \circ e$ is an acceptable numbering of (D', e'). Finally, $\phi : D \to D'$ is a **principal numbering** (of (D', e') w.r.t. (D, e)) if ϕ is a morphism and $\forall \psi \in Mor(D, D') \; \exists \theta \in Mor(D, D) \; \psi = \phi \circ \theta$. Clearly acceptable, relative and principal numberings are onto maps and generalize Gödel-numberings.

Consider now $E = Cont(\omega, \omega^{\perp})$, where ω is given the discrete topology. Clearly, E is the effective domain of the partial functions $E_c = PR$. Set $E^{(n+1)} = Cont(E^{(n)}, E^{(n)})$. We can

then state the main theorem in [48].

Theorem 4.2 *For all $n \geq 1$ one has :*
(1) $C^{(n)} = E_c^{(n)}$
(2) $C^{(n)} \times C^{(n)} \cong C^{(n)}$ *via* $<,>$
(3) $C^{(n.5)} = Cont(E^{(n-1)}, E^{(n)})_c$
(4) $\exists \phi \in C^{(n.5)}$ *principal relative numbering.*

Proof. [48]. □

Properties (1) - (4) are proved by combined induction. (1) and (3) give the continuity of the HPEF. More precisely, (2) at type n+1 is obtained from (1) + (2) at type n. by first showing that, in an effective domain, any effective coding of pairs of algebraic elements by algebraic elements is inherited to a corresponding effective coding at the higher type (the function space). (1) + (3) + (4) give (1) at the higher type. (1) + (2), again, prove (3) and so on, with various combinations, to prove the rest.

Note, now, that once proved that each $C^{(n)}$ is a countable set, with the natural numbering of constructive domains, (P) above amounts to say that each $C^{(n.5)}$ contains a principal numbering. In fact, by generalized Myhill-Shepherdson theorem (see Corollary 2.7), the continuous and computable functions over the computable elements of effective domains coincide with their morphisms, as numbered sets.

There may be direct arguments for proving the arithmetic statement (P), i.e. with no use of the continuous structure. In [48], though, it seemed essential to use finite approximants (algebraic elements) and their effectiveness properties to deal with the functions in the $C^{(n)}$'s, which are infinite objects.

Remark. Topological or limit structures are implicit in the definition of the Kleene-Kreisel countable functionals, when given by the "associates" (see [55]). This is not so for the HPEF, where in the definition, there is no flavour of continuity nor approximation. Though, as soon as we just want to check if the definition makes sense in any type, finite approximants, that is the algebraic elements of suitable domains, come in. Of course, Myhill-Shepherdson theorem and its generalization is the key fact to understand this ; once more, that theorem provides the main justification for Scott's continuity and its deep connections to computability.

5 Invertible terms in λ-calculus

Terms of λ-calculus are defined by variables, formal application and the binding operator λ. The computation (reduction) rules are axiomatized as follows :

$$(\beta) \quad (\lambda x.M)N \geq [N/x]M \quad , \quad \text{for } N \text{ free for } x \text{ in } M, \text{ as usual} ;$$

$$(\eta) \quad \lambda x.Mx \geq M \quad , \quad \text{for } x \text{ not free in } M.$$

The reduction predicate "$\geq$" is reflexive, transitive and substitutive. "$=$" is the least equivalence relation such that $M \geq M' \Rightarrow M = N$; "$=$" also is substitutive. A term is in

(β-)normal form ($\beta\eta$-normal form) if no subterm is the LHS of axiom $(\beta)((\beta)$ or $(\eta))$. A term M **possesses a ($\beta\eta$-) normal form** if, for some M' in $(\beta\eta$)normal form, one has $\lambda\beta(\eta) \vdash M = M'$. Finally, M is **solvable** (or has head normal form), if, for some $P_1, ..., P_q, \vec{x}$ and y, $\lambda\beta \vdash M = \lambda x_1...x_n.yP_1...P_q$, M is **unsolvable**, otherwise. The paradigmatic role of λ-calculus and Combinatory logic, as functional programming languages, is well known since long, see [5].

5.1 The problem

Combinatory Logic and λ-calculus, as type free applicative languages, have a set of terms which forms a monoid with $I \equiv \lambda x.x$ as identity and composition "$\circ$" as application, where $M \circ N = BMN$ for $B = \lambda xyz.x(yz)$. Church proved that the closed terms of $\lambda\beta\eta$-calculus form a recursively presented monoid, with an unsolvable word problem. Note, now, that equations among terms give a standard methods for defining terms, as well as for programs. Equations, though, as well known from algebra, can be more easily solved when working in a group. Consider, for example, the following definition of the (unknown) term $X : M \circ X = X \circ P$. If M, say, possesses an inverse, M^{-1} , an easy application of Curry's fixed point combinator gives X from $X = M^{-1} \circ X \circ P$, since we obtain an ordinary recursive definition of X. It is then an important question, raised by [12], under what conditions a term in the monoid has an inverse.

Definition 5.1 *Let T be (an equational extension of) λ-calculus. M is T-***invertible** *if $\exists N\ T \vdash M \circ N = I$ and $T \vdash N \circ M = I$. (In this case we that M, N are T-***inverse***).*

[18] gives a characterization of invertible terms in $\lambda\beta\eta$-calculus, among those which *possess a normal form* (for $\lambda\beta$-calculus, the group coincides with I). This was done by an (implicit) use of the notion below of tree for terms.

For the purposes of the main theorem of this section, a fundamental definition is required : a term of λ-calculus may be seen as a tree, by taking as root the "head" of the term (e.g. $\lambda x.y$ in $\lambda x.yM$), hereditarily.

Definition 5.2 *(Informal) The* **Böhm tree** *of M is given by :*

$BT(M) = \Omega$ *if M has no head normal form (i.e. M is unsolvable)*

$BT(M) = \lambda x_1...x_n.y$ $\qquad if\ M =_\beta \lambda x_1...x_n.yM_1...M_p$

$BT(M_1) \quad \cdots \quad BT(M_p)$

The definition is informal, since Böhm-trees may be infinite (see [6] for a precise definition). In particular, $BT(M)$ finite and Ω-free iff M has a normal form. Thus, there are plenty of interesting terms with infinite Böhm-trees. Take, say the fixed point combinator $Y = \lambda y.(\lambda x.y(xx))(\lambda x.y(xx))$ or the solution of $X = \lambda z.zX$, and a lot more. Since early work by Wadsworth and Hyland, Böhm-trees have been a basic tool in the comparison of operational

and denotational semantics of λ-calculus (see [6], [45]). Roughly, this is because all terms, as all programs, are finite Böhm-trees "display" their computational (operational) behavior, which may be infinite (see [46] for a discussion).

Böhm-trees may be partially ordered by setting Ω, the undefined tree, as the least one and, then, by setting $BT(M) \subseteq BT(N)$ if $BT(N)$ is obtained from $BT(M)$ by replacing Ω in some leaves of $BT(M)$ by some Böhm-tree. Dezani's result (see theorem 5.4, below) uses a difficult combinatory technique, where the finiteness of the Böhm-trees considered has an essential role. The question remained of characterizing all invertible terms, not necessarily with normal form; equations like the example above are perfectly sound also if $BT(M)$ is infinite, for M could still possess a head normal-form.

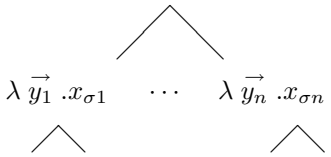
[8] fully answered the general questions, by using results in [18][3]. Our present interest in the methods used in [8] relies on the essential application made of infinitary, continuous structures to solve this typically combinatorial problem. What's nice is that the characterization in [8] confirms the result in [18], since invertible terms turn out to live only among normal forms. To see this, though, one needs to look also at terms with an infinite computational behavior.

5.2 The method

Let $\sigma_1, ..., \sigma_n$ be a permutation of $1, ..., n$ and $\vec{z}$ be a finite string of variables.

Definition 5.3 *(i) The* **hereditary permutators** *(HP) is the set of λ-terms defined by :*

$M \in HP \quad if \quad BT(M) = \lambda z x_1 ... x_n . z$



and each subtree is the tree of a term in HP.

(ii) The **finite hereditary permutators** *(FHP) are the terms in HP with a finite, Ω-free Böhm-tree.*

Theorem 5.4 *Assume that M has a normal-form. Then M is $\beta\eta$-invertible iff $M \in FHP$.*

Proof. [18] (see also [8]). □

Theorem 5.5 *M is $\beta\eta$-invertible iff $M \in FHP$.*

Proof. [8] (see also [6]). □

It is clear that in order to prove theorem 5.5 from theorem 5.4 one only has to show

(O) M is $\beta\eta$-invertible $\Rightarrow$ M has a normal form,

since any M in FHP automatically has a normal form.

[3] A categorical application of invertible terms is given in [11], as these terms are used to characterize the isomorphisms which hold in all Cartesian Closed Categories.

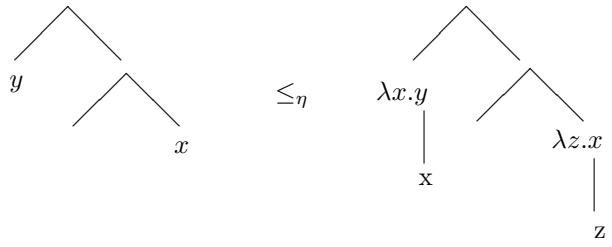
Bergstra and Klop give two arguments for (O), by embedding λ-terms into different continuous models, D_∞ and $P\omega$ (see section 3). Observe first that the valid equalities between λ-terms in $P\omega$ are characterized by the following theorem.

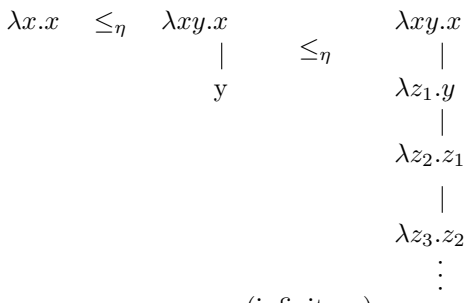
Theorem 5.6 *(Hyland)* $P\omega \models M = N$ *iff* $BT(M) = BT(N)$.

Proof. (see [6]). □

We also need to introduce an interesting extension of $\lambda\beta\eta$-calculus, which may be defined in terms of Böhm-trees. Note first that, once the (possibly infinite) operational behaviour of a λ-term has been displayed by its Böhm-tree, one may also consider infinitary reduction rules (or expansion rules, their inverse operations). In particular, one may η-expand (possibly infinitely many) nodes of a tree. Call "$=_\eta$" such an equality obtained by (possibly infinitely many) η-expansions between trees.

Examples





It is easy to see that the last tree is $BT(J)$, for $J = \lambda xy.x(Jy)$.

Definition 5.7 $H^\star = \{M = N \mid M, N$ *are* λ*-terms and* $BT(M) =_\eta BT(N)\}$

$H^\star$ is a consistent extension of $\lambda\beta\eta$-calculus. It is actually the largest consistent extension of the theory H, where all unsolvable terms are equated ([6]). Thus $H^\star$ is the unique Hilbert-Post completion of H.

Theorem 5.8 *If* M *is* $\beta\eta$*-invertible, then* $M \in HP$.

Proof. (Sketch) Assume that M, N are $\beta\eta$-inverse. Then M, N are also $H^\star$-inverse, since $\lambda\beta\eta \subseteq H^\star$. Write now $M \in_k HP$ if M is in HP up to level k of $BT(M)$ (the root is level 0 of a tree).

If we prove

$$1) \qquad \forall k\ (M\ H^\star\text{-invertible} \Rightarrow M \in_k HP),$$

we are done, by the inductive definition of HP.

1) above holds for $k = 1$, by (a lot of) combinatory work on $BT(M)$, up to the first two levels (see [18], [8] or [6]).

Assume now (1) for $k > 1$. Then, by some further non-obvious work, prove that, if M, N $H^\star$-inverse, then $N_i, M_{\sigma i}$ are $H^\star$-inverse.

Thus, by induction hypothesis, $N_i, M_{\sigma i} \in_k HP$ and, clearly, $M, N \in_{k+1} HP$. □

The key idea is that, by the infinitary character of the provable equalities, $H^\star$ allows to compare terms with infinite Böhm-trees. Roughly, it is like looking at full decimal expansions of rational numbers, without restricting the attention to numbers with finite expansions (normal forms). The proof, though, works, at finite levels, exactly as for the finitary result and it is then analytically extended to the infinite expansions.

Church original work on λ-conversion used a more restrictive notion of λ-term. Namely, λI-terms are defined similarly to λ-terms, provided that $\lambda x.M$ is a λI-term iff x occurs free in M. It is easy to see that a λI-term M has a normal form iff all of its subterms have a normal form. By looking at the definition of HP, one may work out the following fact.

Theorem 5.9 *If $M \in HP$, then $BT(M)$ is Ω-free and there exists a λI-term M' such that $BT(M') = BT(M)$.*

Let's now follow stepwise Bergstra and Klop's argument for

(O) If M is $\beta\eta$-invertible, then M has a normal form. Assume that M, N are $\beta\eta$-inverse and that M has no normal form. Then

(1) There exist λI-terms M', N' s.t. $BT(M') = BT(M)$ and $BT(N') = BT(N)$, by theorem 5.9.

(2) $BT(M')$ is infinite, by assumption and theorem 5.8, and, hence M' has no normal form.

(3) $M' \circ N'$ is a λI-term, for $M' \circ N' \equiv BMN$ where $B \equiv \lambda xyz.x(yz)$, and, hence, it has no normal form, by (2).

(4) $P\omega \models M' = M$ and $P\omega \vdash N' = N$. by (1) and theorem 5.6.

(5) $P\omega \models M' \circ N' = M \circ N$. by (4).

(6) $BT(M' \circ N') = BT(M \circ N)$, by (5) and theorem 5.6, again.

(7) $BT(M \circ N)$ is infinite, by (3) and (6), and, hence , $M \circ N$ has no normal form.

The conclusion in (7) is clearly impossible, for $\lambda\beta\eta \models M \circ N = I$.

Comment 5.10 *The role of $P\omega$, as continuous topological model should be clear. It essentially allows to go from (1) to (6), by using theorem 5.6. Since BT-equal terms are the same object in $P\omega$ by theorem 5.6, one has (4) and, hence, (5) and (6) trivially follow.*

It is like looking at a property of natural numbers by studying them as elements of a more structured continuum they live in, the real or complex field, say. Where does continuity come in, in this case ? The proof of theorem 5.6 heavily relies on continuity properties of $P\omega$. To see this more closely, one may consult [6] or look at results which distillate the notion of "approximable application" in [45]. This notion relates, in a general setting, the interpretation of formal application to the limit structure of models. Can one directly go from (1) to (6) by working over Böhm-trees with some smart combinatory technique ? This is done in [6], but the continuity argument is important as well. As a matter of fact one has to take the *completion* of the set of Böhm-trees, partially ordered as after definition 5.2, and heavily work with approximation and Scott continuity over this domain.

A final remark on cardinality. It is possible to take a countable model instead of $P\omega$, in the argument above, since $RE \subseteq P\omega$, the set of recursively enumerable sets, is an equationally equivalent sub-model of $P\omega$. That is, by the same notions of application and abstraction, RE yields a model which satisfies the same equalities as $P\omega$. More generally, for any infinite cardinal α, there is an applicative structure of cardinal α which yields lots of different λ-models, including one equationally equivalent to $P\omega$ (see [45]). The use of RE, though, instead of $P\omega$ would be like looking at rational numbers (or complex numbers with rational coordinates), for the analysis of the behaviour of continuous functions or predicates over the real (or complex) numbers. Those countable subsets are dense and they perfectly determine continuous functions or predicates on the latter sets. Similarly, RE is dense in $P\omega$ w.r.t. the Scott topology (and contains all algebraic elements).

It should be clear that this informal survey is far away from being exhaustive as far as the use of "analytic" methods in Computer Science is concerned. We hope, though, that the examples mentioned provide some sufficiently clear analyses of combinatorial facts within continuous/topological frameworks even in Theory of Computations.

Bibliographie

[1] S. Abramsky, A. Jung , *"Domain Theory"*, in Abramsky S. et al. (eds.) Handbook of logic in computer science, Vol. I and 2, Oxford : Clarendon Press, 1992.

[2] L. Aceto, V. Bjorn, G. Longo (eds.) *The difference between Turing Computability and Concurrent Systems*, special issue Mathematical Structures in Computer Science, Cambridge U.P., to appear, 2002.

[3] R. Amadio, P.-L. Curien, *Domains and lambda-calculi*, Cambridge Univ. Press, 1998.

[4] A. Asperti, G. Longo, *Categories, Types and Structures*, M.I.T. Press, 1991.

[5] J. Backus, *Can Programming be liberated from the Von Neumann style ? A functional style and its Algebra of Programs*, Communications of ACM, 21,8, 1978.

[6] H.P. Barendregt, *The lambda calculus : its syntax and semantics*, North-Holland, Amsterdam (revised edition), 1984.

[7] H.P. Barendregt, G. Longo, *"Recursion theoretic operators and morphisms of numbered sets"*, Fundamenta Mathematicae, (119) :49–62,1982.

[8] J. Bergstra, J.W. Klop, *"Invertible terms in the lambda-calculus"*, Theor. Comp. Sci. ll, 19-37, 1980.

[9] G. Berry, P.L. Curien, *"Sequential Algorithms on Concrete Data Structures"*, Theor. Comp. Sci. 20, (265-321), 1982.

[10] L. Birkedal, J. van Oosten, G. Rosolini, D.S. Scott (eds.) *Realizability Semantics and Applications*, special issue Mathematical Structures in Computer Science, vol. 12, n. 3, 2002.

[11] K. Bruce, R. Di Cosmo, G. Longo, *"Provable isomorphisms of types"*, Mathematical Structures in Computer Science, 2(2) :231–247, 1992.

[12] H. B. Curry, R. Feys, *Combinatory Logic*. North-Holland, Amsterdam, 1958.

[13] G. Castagna, *Object-Oriented Programming : A Unified Foundation*, Birkäuser, 1997.

[14] A. Church, *"A set of partulates for the Foundation of Logic"* Annals of Math. XXXIII (348-349) and XXXIV (839-864), 1932.

[15] A. Church, *The Calculi of Lambda Conversion*, Princeton Univ. Press, reprinted 1963 by University Microfilms Inc., Ann Arbor, Michigan, U.S.A, 1941.

[16] H.B. Curry, *"Grundlagen der kombinatorischen Logik"*, Amer. J. of Math., 52, 1930.

[17] F. De Jaeger, *"An approach to Effective Functionals on Real Numbers via Filter Spaces"*, proceedings of Summer Conference Series in Topology and Applications, 2002.

[18] M. Dezani, *"Characterization of normal forms possessing inverse in the $\lambda\beta\eta$-calculus"*, Theor. Comp. Sci. 2, 323-337, 1976.

[19] A. Edalat, *"Dynamical systems, Measures and Fractals via Domain Theory"*, Information and Computation 120(1), (1995), 32-48.

[20] A. Edalat, *"Domain Theory and Integration"* Theoretical Computer Science 151, (1995), 163-193.

[21] A. Edalat, R. Heckmann, *"A Computational Model for Metric Spaces"*, Theoretical Computer Science 193, (1998) 53-73.

[22] A. Edalat, *"When Scott is Weak on the Top"*, Mathematical Structures in Computer Science 7, (1997), 401-417.

[23] A. Edalat, *"Domains for Computation in Mathematics, Physics and Exact Real Arithmetic"*, Bulletin of Symbolic Logic, 3(4), (1997), 401-452.

[24] M. Escardo, *"PCF extended with real numbers"*, Theoretical Computer Science, vol. 162, 79-115, 1996.

[25] Yu. L. Ershov, *"Model C of partial continuous functionals"*, In Logic Coll. 76, eds. R. Gandy, M. Hyland. North-Holland, Amsterdam, 1976.

[26] G. Frege, *The Foundations of Arithmetic*, 1884 (Engl. transl. Evanston, 1980.)

[27] G. Gierz et al., *A compendium of Continuous lattices*, Springer-Verlag, Berlin, 1980.

[28] J.Y. Girard, *"The system F of variable types, fifteen years later"* Theoretical Comp. Sci., 45 (159-192), 1986.

[29] J.-Y. Girard, *"Linear Logic"*, Theoretical Comp. Sci., 50 (1-102), 1987.

[30] J.Y. Girard, Y. Lafont, P. Taylor, *Proofs and Types*, Cambridge U. Press, 1989.

[31] J.Y. Girard, *"Locus Solum"*, Mathematical Structures in Computer Science, vol.11, n.3, 2001.

[32] K. Gödel, *"Uber formal unentscheidbare Satze der Principia mathematica und verwandter Systeme I"*, Monastshefte f. Mathematik und Physik 38, 1931.

[33] H. Goldfarb, *Jacques Herbrand : logical writings*, 1987.

[34] E. Goubault (ed.), *Geometry in Concurrency*, special issue, Mathematical Structures in Computer Science, Cambridge U.P., vol.10, n.4, 2000.

[35] A. Heyting, *Constructivity in Mathematics*. North-Holland, 1959.

[36] D. Hilbert, *Les fondements de la géométrie*, 1899 (trad. fran., Dunod, 1971).

[37] M. Hyland, *"Filter Space and Continuous Functionals"*, Annals of Mathematical Logic 16, (101-143), 1977.

[38] M. Hyland, *"The effective Topos"*, in The Brouwer Symposium, (Troelstra, Van Dalen eds.) North-Holland, 1982.

[39] M. Hyland, *"A small complete category"*, Lecture delivered at the Conference Church's Thesis after 50 years, Zeiss (NL), June 1986 (Ann. Pure Appl. Logic, 1987).

[40] M. Hyland, A. Pitts, *"The Theory of Constructions : categorical semantics and topos theoretic models"*, Categories in Comp. Sci. and Logic, Boulder (AMS notes), 1987.

[41] S. C. Kleene, *"Lambda definability and recursiveness"*, Duke Math. J., 2, (pp. 340-353), 1936.

[42] S.C. Kleene, *"Recursive functionals and quantifiers of finite type I"*, TRANS AMS, vol. 91, 1959 (and II, vol. 95, 1963).

[43] A. Joyal, M. Nielsen, G. Winskel, *"Bisimulation from Open Maps"*, Information and Computation 127(2) : 164-185, 1996.

[44] F.W. Lawvere, *"Variable quantities and variable structures in topoi"*, in Algebra Topology and Category Theory : a collection of papers in honor of Samuel Eilenberg, A. Heller and M. Tierney (eds.), Academic Press, (101-131), 1976.

[45] G. Longo, *"Set-theoretical models of λ-Calculus : theories, expansions, isomorphisms"*, Ann. Pure Applied Logic, 24, 153-188, 1983.

[46] G. Longo, *"Continuous structures and Analytic Methods in Computer Science"*, in **Ninth Colloquium on Trees in Algebra and Programming** (CAAP 84), pages 1–22 (Courcelle, editor). Cambridge University Press, 1984.

[47] G. Longo, *"Space and Time in the foundation of Mathematics, or some challenges in the interaction with other sciences"*, invited lecture at the First AMS/SMF meeting, Lyon, July 2001 (to appear).

[48] G. Longo, E. Moggi, *"The hereditary partial recursive functionals and recursion theory in higher types"*, Journal of Symbolic Logic, 49(4) :1319 – 1332, 1984.

[49] G. Longo, E. Moggi, *"A category-theoretic characterization of functional completeness"*, Theor. Comp. Sci. vol. 70, 2, (pp. 193-211), 1990.

[50] G. Longo, E. Moggi, *"Constructive Natural Deduction and its omega-Set Interpretation"*, Mathematical Structures in Computer Science,vol.1, n.2, 1991.

[51] S. Mac Lane, I. Moerdijk *Sheaves in Geometry and Logic*. Springer Verlag, Berlin, 1992.

[52] J. Myhill, J. C. Shepherdson, *"Effective operations on partial recursive functions"*, Zeit. f. Math. Logik und Grund. der Math.,vol. 1, 1955.

[53] Y. Moschovakis, *"Recursive metric spaces"*, Fundam. Mathematicae, LV, 1964.

[54] M. Nivat, *"Infinite words, infinite trees, infinite computations"* , Math. Cent. Tracts 109, 1-52, 1979.

[55] D. Normann, *Recursion on the countable functionals* LNM 811, Springer-Verlag, Berlin, 1980.

[56] A. Pitts, *"Polymorphism is Set Theoretic, constructively"*, Symposium on Category Theory and Comp. Sci., SLNCS 283 (Pitt et al. eds), Edinburgh, 1987.

[57] H. Rogers, *Theory of Recursive Functions and Effective Computability*, 1967.

[58] H. Rice, *"Classes of recursively enumerable sets and their decision problem"*, Trans. AMS, vol. 74, 1953.

[59] B. Riemann, *"On the hypothesis which lie at the basis of geometry"*, 1854 (English transl. by W. Clifford, Nature, 1873).

[60] D. S. Scott, *"Outline of a Mathematical Theory of computation"*, 4th Ann. Princeton conf. on Info. Sci. and Syst., 1970.

[61] D.S. Scott, *"Continuous lattices"*, In Toposes, Algebraic Geometry and Logic. LNM 274, Springer-Verlag, Berlin, 1972.

[Scott76] D.S. Scott, *"Data types as lattices"*, SIAM J. of Comput., 5, 522-587, 1976.

[62] M. Smyth, G. Plotkin, *"The category-theoretic solution of recursive domain equations"*, SIAM Journal of Computing 11, 1982.

[63] J. Tappenden, *"Geometry and generality in Frege's philosophy of Arithmetic"*, Synthese, n. 3, vol. 102, March 1995.

[64] A. Turing, *"On computable numbers with an application to the entscheidungsproblem"*, PROC. London Math. soc. 42, 1936.

[65] H. Weyl, *Philosophy of Mathematics and of Natural Sciences*, 1927 (Engl. transl., Princeton University Press, Princeton, New Jersey, 1949).

[66] G. Winskel, M. Nielsen, *"Models of Concurrency"*, Handbook of Logic and the Foundations of Computer Science, vol IV, Abramsky et al. editors, Oxford University Press, 1995.

Géométrie et syntaxe : les ensembles semi-algébriques et leurs généralisations

Marie-Françoise Roy[1]
IRMAR, CNRS, Université de Rennes I,
Campus de Beaulieu 35042 Rennes cedex France

1 Sturm-Sylvester

Avant de parler des ensembles semi-algébriques, il est utile d'étudier la question suivante : peut-on compter exactement les racines réelles d'un polynôme ? Une méthode dûe à Sturm [10, 8, 5] permet de le faire par un calcul qui se déroule tout entier dans le corps engendré par les coefficients du polynôme. La base de ce calcul est la division euclidienne des polynômes.

Plus précisément, soient P et Q des polynômes en une variable à coefficients réels. La suite de Sturm de P et Q est la suite de polynômes $P_0, ..., P_k$ définie par $P_0 = P, P_1 = P'Q, P_i = P_{i-1}C_i, -P_{i-2}$ avec C_i polynôme à coefficients réels $\deg(P_{i+1}) < \deg(P_i)$ pour $i = 2, ..., k$ et P_k est un plus grand commmun diviseur de P et $P'Q$.

Si a réel n'est pas une racine de P et $P_0, ..., P_k$ est la suite de Sturm de P et Q, on définit $v(P, Q; a)$ comme le nombre de changements de signes dans $[P_0(a), ..., P_k(a)]$.

On alors (Théorème de Sylvester) [10, 2, 1] : La différence entre le nombre de racines de P dans l'intervalle $]a, b[$ pour lesquelles Q est positif et le nombre de racines de P dans l'intervalle $]a, b[$ pour lesquelles Q est négatif, est égale à $v(P, Q; a) - v(P, Q; b)$.

Si $Q = 1$, on peut compter ainsi les racines de P sur la droite toute entière, puisque les variations de signes dans la suite de Sturm à l'infini ne dépend que des signes des coefficients dominants et de la parité des degré des polynômes de la suite. On peut également compter les racines de P rendant Q positif sur la droite toute entière, en comptant, si P et Q sont premiers entre eux, le nombre des racines de P (par la suite de Sturm de P et 1), puis la différence entre le nombre de racines de P pour lesquelles Q est positif et le nombre de racines de P pour lesquelles Q est négatif (par la suite de Sturm de P et Q). Il est aussi possible de décider par une méthode algorithmique si un polynôme P a une racine vérifiant

[1]Marie-Francoise.Roy @univ-rennes1.fr

des contraintes de signes supplémentaires exprimées par des conditions de signes portant sur d'autres polynômes $Q_1, ..., Q_s$: on utilise le théorème de Sylvester de nombreuses fois pour le polynôme P et des produits des polynômes Q_i [1].

Finalement, on ne sait pas exprimer les racines réelles d'un polynôme par des formules, puisqu'il est impossible de résoudre par radicaux à partir du degré 5, mais on sait bien compter le nombre des racines réelles, y compris avec contraintes polynomiales [1].

Si un polynôme a des coefficients entiers, le calcul de la suite de Sturm ne fait intervenir que des nombres rationnels. L'algorithme fonctionne en fait pour un corps réel clos quelconque (un corps réel clos est un corps ordonné où tout nombre positif est un carré et où tout polynôme de degré impair a une racine).

2 Les ensembles semi-algébriques

Une classe très naturelle d'objets géométriques est celle des ensembles semi-algébriques. Les ensembles semi-algébriques de R^n sont engendrés par combinaison booléenne à partir des ensembles définis par un égalité ou une inégalité polynomiale.

Les propriétés géométriques des ensembles semi-algébriques sont tout à fait remarquables. Par exemple [2, 3]

a) les ensembles semi-algébriques forment une classe stable par projection,

b) l'adhérence, l'intérieur d'un ensemble semi-algébrique est semi-algébrique,

c) les ensembles semi-algébriques ont un nombre fini de composantes connexes, elles-mêmes semi-algébriques,

d) tout ensemble semi-algébrique fermé borné est homéomorphe a un complexe simplicial fini,

e) tout ensemble semi-algébrique peut être stratifié, c'est-à-dire décomposé en un nombre fini de variétés lisses de différentes dimensions, qui sont elles-mêmes semi-algébriques,

f) (curve selection lemma) tout ensemble semi-algébrique est tel que de tout point de son adhérence on peut faire partir un chemin continu contenu dans l'ensemble sauf en son point de départ.

Pour résumer : les ensembles semi-algébriques présentent toutes les propriétés de finitude qu'on peut espérer, et un comportement tout à fait modéré.

Ces ensembles ont par ailleurs une description syntaxique bien précise. Précisons les définitions. Une formule atomique du langage des corps ordonnés est une formule $P > 0$ ou $P = 0$ où P est un polynôme en plusieurs variables à coefficients réels. Les formules sans quantificateurs du langage des corps ordonnés sont obtenues à partir des atomes en utilisant les connecteurs logiques " et ", " ou " et " non ". Les ensembles semi-algébriques sont donc les ensembles décrits par des formules sans quantificateurs du langage des corps ordonnés.

Les formules du premier ordre du langage des corps ordonnés sont plus générales : elles sont construites à l'aide des atomes et des connecteurs logiques précédents en utilisant les quantifications existentielles et universelles portant sur les éléments du corps. Cette restriction est importante : on ne peut donc pas quantifier sur les entiers, par exemple, mais seulement sur les réels. Suivant la terminologie de Tarski [11, 8], un ensemble définissable

de nombres réels est un ensemble décrit par une formule du premier ordre du langage des corps ordonnés.

Certaines propriétés géométriques des ensembles semi-algébriques ont une traduction syntaxique. La stabilité des ensembles semi-algébriques par projection (propriété a)) n'est rien d'autre qu'une reformulation géométrique du fameux principe de Tarski-Seidenberg [11] qui, sous sa forme logique, affirme que " la théorie des corps réels clos admet l'élimination des quantificateurs dans le langage des corps ordonnés". En effet, on peut éliminer les quantificateurs un par un, il suffit d'éliminer les quantificateurs existentiels puisque " pour tout $xP(x)$" n'est autre que " non il existe $xnon\ P(x)$" et éliminer un quantificateur existentiel n'est rien d'autre que dire que la projection d'un ensemble semi-algébrique est semi-algébrique [2]. Autrement dit, les ensembles définissables de nombres réels sont les ensembles semi-algébriques.

On peut penser, à lire les paragraphes précédents, que la formulation logique est redondante et qu'il suffit de s'en tenir au point de vue géométrique. Ce n'est pas si simple. L'attention portée à la syntaxe est utile. Pour montrer que l'adhérence d'un ensemble semi-algébrique est semi-algébrique (propriété b) il est plus commode d'écrire la description syntaxique de l'adhérence par une formule du premier ordre que de faire intervenir des projections et des complémentaires en cascade [2].

La démonstration par Tarski de son théorème d'élimination des quantificateurs repose sur une induction sur le nombre de variables [12, 3, 1]. Si on part de n variables, on considère $n-1$ d'entre-elles comme des paramètres et la dernière comme variable principale. En utilisant le théorème de Sylvester, on obtient, grâce aux coefficients dominants de diverses suites de Stum-Sylvester, des conditions de signes portant sur des polynômes en $n-1$ variables qui contrôlent l'existence de racines réelles de polynômes en n variables vérifiant certaines contraintes de signes. Le fait que le comptage des racines avec contraintes de signes se fait par des calculs de division euclidienne est ici essentiel. Cette particularité implique aussi que le processus d'élmimination des quantificateurs est effectif, et que l'algébre et la géométrie élémentaires sont décidables [12]. Des études quantitatives plus précises sur la complexité de l'élimination des quantificateurs restent d'actualité [1].

Une conséquence immédiate du théorème de Tarksi est que tout ensemble définissable de la droite réelle est formé d'un nombre fini de points et d'intervalles. Dès qu'on ne respecte pas la syntaxe, on perd cette propriété de finitude : le sous-ensemble des nombres réels formé de l'ensemble des nombres entiers n'est pas un ensemble définissable de nombres réels, puisqu'il n'est pas formé d'un nombre fini de points et d'intervalles. Sa définition par une formule fait d'ailleurs intervenir une quantification portant sur les entiers, ce qui n'est pas permis pour une formule du premier ordre.

Les autres propriétés c), d), e) et f) n'ont pas de pendant logique particulier, elles sont aussi montrées par induction sur le nombre de variables [2]. Le cas de base est vrai pour les ensembles semi-algébriques de la droite qui sont formés d'un ensemble fini de points et d'intervalles et pour lesquelles les propriétés c) d) et e) sont immédiates.

3 o-minimalité

Cette théorie propose une axiomatique des ensembles vérifiant les propriétés similaires aux semi-algébriques [6] (voir aussi [4]).

Une structure o-minimale est composée pour chaque entier n d'une classe S_n de parties de R^n avec les propriétés suivantes

1) S_n contient les ensembles définis par un égalité ou une inégalité polynomiale,

2) S_n est clos par opération booléenne,

3) la projection d'un ensemble de S_n sur R^{n-1} est un ensemble de S_{n-1} ,

4) les éléments de S_1 sont les unions finies de points et d'intervalles.

L'axiome 4) est fondamental, les axiomes 1), 2) et 3) expriment simplement que les ensembles de la structure o-minimale sont les ensembles définissables par les formules d'une théorie du premier ordre exprimée dans le langage des corps ordonnés, où on dispose de symboles de fonctions éventuels en plus de $+$ et $\times$.

Il est clair que les ensembles semi-algébriques forment une structure o-minimale.

Les propriétés géométriques des ensembles d'une structure o-minimale sont tout à fait semblables à celles des ensembles semi-algébriques

a) les ensembles d'une structure o-minimale forment une classe stable par projection,

b) l'adhérence, l'intérieur d'un ensemble d'une structure o-minimale appartient à cette structure,

c) les ensembles d'une structure o-minimale ont un nombre fini de composantes connexes, qui appartiennent à la structure,

d) tout ensemble d'une structure o-minimale fermé borné est homéomorphe a un complexe simplicial fini,

e) tout ensemble d'une structure o-minimale peut être stratifié, c'est-à-dire décomposé en un nombre fini de variétés lisses de différentes dimensions, qui appartiennent à la structure,

f) (curve selection lemma) tout ensemble d'une structure o-minimale est tel que de tout point de son adhérence on peut faire partir un chemin continu contenu dans l'ensemble sauf en son point de départ.

Un exemple typique d'une fonction dont le graphe ne peut pas faire partie d'une structure o-minimale est la fonction $\sin(1/x)$ définie sur la droite réelle. Ce graphe a un comportement fort peu modéré. Tout point de la forme $(0, a)$, $a \in [0, 1]$ appartient à son adhérence et pourtant il n'existe aucun chemin continu partant de $(0, a)$ qui soit contenu tout entier à l'intérieur du graphe de $\sin(1/x)$ sauf en son point de départ. Le point f) de la liste ci-dessus est donc contredit.

Ici a) et b) résultent directement des axiomes, puisque la projection et l'adhérence d'un ensemble définissable est définissable. Les démonstrations des autres résultats c) , d), e) et f) s'inspirent de près des preuves connues pour les ensembles semi-algébriques et sont basées sur un procédé inductif en la dimension de l'espace ambiant, dont le cas de base est réglé par l'axiome 4).

Là encore : les ensembles définissables d'une structure o-minimale présentent toutes les propriétés de finitude qu'on peut espérer, et un comportement tout à fait modéré.

Notons que l'accent mis dans le cas des ensembles semi-algébriques sur l'élimination des quantificateurs n'a plus cours, puisque la stabilité par projection fait partie des axiomes. Il est en général faux que les ensembles définissables d'une structure o-minimale soient décrits pas des formules sans quantificateurs. De même les résultats d'effectivité et de décidabilité qui résultaient du théorème de Sturm et de Sylvester ne se généralisent pas.

Outre le cas des ensembles semi-algébriques, de nombreux exemples de structures o-minimales d'une grande importance mathématique ont été découverts et sont étudiés : ensembles sous-analytiques, ensembles définissables avec l'exponentielle, ensembles semi-pfaffiens [7].

L'étude de la géométrie o-minimale se développe actuellement dans deux directions : l'étude systématique des conséquences des axiomes et la découverte de situations mathématiquement significatives où les axiomes sont vérifiés.

La géométrie algébrique et analytique réelle reste donc un lieu de contact et de dialogue entre la logique et la géométrie, au point de créer un micro-climat surprenant : la distinction entre logicien et géomètre n'a plus cours dans la petite communauté des chercheurs actifs de ce domaine.

Bibliographie

[1] S. Basu, R. Pollack, M.-F. Roy, *Algorithms in real algebraic geometry*, Springer Verlag, 2003.

[2] J.Bochnak, M. Coste, M.-F. Roy, *Géométrie algébrique réelle*, Springer-Verlag, 1987. *Real algebraic geometry*, Springer-Verlag, 1998.

[3] M. Coste, *An introduction to semi-algebraic geometry*, Dip. Mat. Univ. Pisa, Dottorato di Ricerca in Matematica, Istituti Editoriali e Poligrafici Internazionali, Pisa, 2000.

[4] M. Coste, *An introduction to o-minimal geometry*, Dip. Mat. Univ. Pisa, Dottorato di Ricerca in Matematica, Istituti Editoriali e Poligrafici Internazionali, Pisa, 2000.

[5] *Histoire d'algorithmes*, Collection Regards sur la science, Belin, 1994.

[6] L.van den Dries, *Tame topology and o-minnimal structures*, London Math. Soc. Lecture Notes 248. Cambridge Univ. Press, 1998.

[7] L.van den Dries, C. Miller, *geometric categories and o-minimal structures*, Duke Math. J. 84, 497-540, 1996.

[8] H. Sinaceur, *Corps et modèles, essai sur l'histoire de l'algèbre réelle*, Vrin, 1991.

[9] C. Sturm, *Mémoire sur la résolution des équations numériques*, Inst. France Sc.Math. Phys. 6, 1835.

[10] J. J. Sylvester, *On a theory of syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's function*, Trans. Roy. Soc. London, 1853.

[11] A. Tarski, *Sur les ensembles définissables de nombres réels*, Fund. Math. 17, 210-239, 1931.

[12] A. Tarski, *A Decision method for elementary algebra and geometry*, University of California Press, 1951.

A propos de l'imbrication des concepts analytiques et géométriques : le livre de Vitushkin sur la complexité du problème des tables

Martin Zerner
Equipe REHSEIS
(UMR 7596, CNRS et Université Paris 7)

1 Introduction

L'objet de cette conférence est un livre qui m'avait conseillé Vladimir Arnold dans ma première rencontre avec lui lors de son premier voyage en France en 1965 c'est celui de Vitushkin sur la complexité du problème de tabulation [11].

Le livre de Vitushkin [11] est très ... complexe et on pourrait l'étudier de beaucoup de points de vue. On pourrait par exemple essayer de comprendre son rapport avec la réalité des pratiques et des recherches en programmation en Union Soviétique à la même époque. On pourrait aussi le prendre pour base d'une étude de cas sur le style soviétique en mathématique. Je prends le mot style dans un sens précis : un système de formes caractérisant un auteur, une école ou une époque. Il faudrait alors étudier les notations (signalons au passage qu'elles sont foisonnantes), l'ordonnancement des lemmes, propositions et théorèmes, le style de l'écriture en langue naturelle, etc. Et naturellement, on pourrait faire une étude classique d'histoire "interne" des mathématiques en le situant dans l'évolution de la discipline, en particulier par rapport aux recherches sur le treizième problème de Hilbert.

Je me limiterai à autre chose. Je chercherai à expliquer le rôle central que joue un objet mathématique, la variation d'un ensemble, de nature géométrique selon moi, inattendu dans ce contexte et à montrer la diversité des domaines mathématiques impliqués. Je parlerai donc en mathématicien et non en historien des mathématiques, en négligeant les précautions

d'usage. Je ne conserverai pas toujours les notations et la terminologie de l'auteur et je simplifierai parfois outrageusement.

Une partie au moins des résultats de Vitushkin ont été obtenus depuis par d'autres méthodes [3] et [4], [13] et [14]. Je n'en pense pas moins que celle qu'il a employée est intéressante en soi. Si j'arrive à en convaincre quelqu'un suffisamment pour qu'il ait envie de lire l'original, j'en serai ravi. Je dois cependant le prévenir d'un certain nombre de difficultés. D'abord, il vaut mieux savoir lire le russe mathématique. Il existe bien une traduction anglaise sous un titre trompeur [12]. Ce qu'on y comprend le plus vite, c'est que le traducteur n'a strictement rien compris à ce qu'il traduisait. J'ai déjà dit que le livre était compliqué, il y a beaucoup d'idées, il y a aussi beaucoup de calculs qui ne demandent certes que les connaissances de la première année d'université mais où les astuces ne sont pas toujours indiquées. Pour aggraver les choses, il y a un bon nombre de coquilles dans les formules. On ne peut pas reprocher à Vitushkin de ne pas s'être relu puisqu'il est aveugle. Mais quelqu'un aurait bien dû le faire sérieusement. Voici quelques indications plus positives. Pour qui ne lit pas le russe, j'ai écrit en français un article où une partie des idées est reprise [15]. Pour qui le lit, un livre ultérieur explique la définition et les propriétés de la variation, des applications à l'analyse et les plus simples de celles que Vitushkin fait au problème des tables [2]. Ajoutons qu'une bonne partie des idées et des résultats se trouvaient déjà dans un livre antérieur du même auteur [10].

Je procéderai en trois parties. Dans la première j'exposerai le problème des tables et la façon dont il conduit à un problème d'approximation dans un espace de Banach, en l'espèce un espace de fonctions muni de la norme du sup. Dans la deuxième, je décrirai la variation d'un ensemble fermé et ses propriétés. On semblera être loin du problème des tables puisque tout se déroulera dans des espaces de dimension finie. Ce sera l'objet de la troisième partie de donner quelques indications sur la façon d'y revenir.

2 Le problème des tables et sa transformation

L'ouvrage paraît donc en 1959. Les ordinateurs, d'ailleurs encore peu répandus, ont des performances qui sont aujourd'hui celles des calculatrices de poche et l'électronique soviétique est en retard sur l'américaine. Qui s'en sert est amené à de délicats compromis entre le volume de mémoire disponible et le temps de calcul. C'est dans ce contexte que Vitushkin s'attaque au problème des tables de fonctions.

Ce problème n'est pas nouveau. On sait bien que les tables de cordes (zij) remontent à l'antiquité et au début du septième siècle, Brahmagupta les lisait déjà par interpolation quadratique de Lagrange [7]. Mais ici, il s'agit de tabuler toutes les fonctions appartenant à une certaine classe.

Pour démontrer des résultats, il faut donner une définition formelle du problème. Soit F une classe compacte de fonctions continues sur un ensemble compact G (la norme est toujours celle du sup). Une table de F est la donnée :

- de la table elle-même, c'est-à-dire qu'à chaque fonction f appartenant à F, on fait correspondre p nombres réels y_j ;
- d'un algorithme de déchiffrement, qui à chaque valeur de x dans G fasse correspondre

une valeur $F_x(y)$ qui approche $f(x)$ à ε près :

$$|f(x) - F_x(y)| \leq \varepsilon.$$

Aucune condition n'est imposée à la table elle-même, ce peut être n'importe quelle application de F dans $\mathfrak{R}^p$. Dans les exemples traités, il s'agit de valeurs ponctuelles des fonctions et de leurs dérivées et de coefficients de développements en série (Laurent, Fourier, Tchebychev). Les fonctions F_x doivent pouvoir être calculées par les opérations numériques de base : les quatre opérations de l'arithmétique et les comparaisons de réels avec saut. C'est ce qu'on fait quand on programme en Fortran. Ces fonctions sont donc des applications définies par des inégalités polynomiales et des fonctions rationnelles. Vitushkin les appelle rationnelles par morceaux et nous appellerons leurs images ensembles rationnels par morceaux (Vitushkin dit surfaces). Rien ne leur est imposé en général concernant la dépendance en x. Une attention spéciale est apportée au cas polynomial :

$$F_x(y) = \sum_k f_k(x) y^k$$

où la somme est finie, $k = (k_1, ..., k_p)$, et les f_i sont quelconques.

Prenons l'exemple des tables usuelles (sinus, logarithmes, ...) et de leur lecture sans interpolation sur un intervalle $[a, b]$. On se donne un nombre n et on pose $h = (b - a)/n$. A la fonction f on fait correspondre les nombres $y_j = f(a + jh)$ et au nombre x de l'intervalle l'indice $j(x) = [(x - a)/n]$. La valeur lue sur la table est alors $y_{j(x)}(x)$. C'est un algorithme polynomial de degré un. Les fonctions F_x s'écrivent :

$$F_x(y) = \sum_k f_k(x) y_k \quad \text{avec} \quad f_k(x) = \chi_k \delta^k_{j(x)}$$

où χ_k est la fonction caractéristique de l'intervalle $[a+kh, a+(k+1)h[$. On notera au passage qu'elle n'est pas continue.

On constate que la définition donnée est extrêmement générale, ce qui donne une très grande puissance aux minorations de complexité obtenues. J'illustrerai la difficulté de ces minorations par une anecdote. Après m'être convaincu de leur justesse, j'ai téléphoné à Jacques-Louis Lions pour lui expliquer ce dont il s'agissait. Il m'a répondu : "Je te crois. Mais si quelqu'un m'avait proposé de chercher de telles inégalités, je lui aurais répondu que c'était impossible."

Les majorations de complexité s'obtiennent en construisant effectivement des tables pour les classes de fonctions les plus usuelles en analyse. Elles sont assez proches des minorations. Je n'en dirai pas plus à leur sujet.

J'ai parlé jusqu'ici de la complexité d'une table. Il y en a deux, l'une qui caractérise le volume en mémoire et l'autre la complexité de l'algorithme de déchiffrement.

La première semble être la dimension p de l'espace de la table. Ce serait oublier la précision avec laquelle les nombres y_j sont stockés. Cependant la caractéristique adéquate avait déjà été introduite par Kolmogorov, c'est l'ε-entropie.

Il faut rappeler les définitions pour l'intelligence de la suite. Soit $M_\varepsilon(F)$ le plus petit nombre d'ouverts de diamètre 2ε au plus permettant de recouvrir F. On appelle ε-entropie (absolue) de F le nombre $H_\varepsilon(F) = log_2 M_\varepsilon(F)$. Soit d'autre part $N_\varepsilon(F)$ le plus grand nombre

de points deux à deux distants d'au moins ε qu'on puisse trouver dans F le nombre $C_\varepsilon(F) = log_2 N_\varepsilon(F)$ est appelé ε-capacité de F. On a la double inégalité facile mais essentielle :

$$C_{2\varepsilon}(F) \leq H_\varepsilon(F) \leq C_\varepsilon(F).$$

Les logarithmes (en base 2) sont là pour des raisons historiques, pour se ramener à des quantités d'information (mesurées en bits). Techniquement, ils ne font guère que compliquer la situation. Les évaluations d'ε-entropie étaient connues, Vitushkin les améliore.

La complexité d'un algorithme de déchiffrement est caractérisée par trois nombres : p que nous connaissons déjà, le degré k des numérateurs et dénominateurs des fractions rationnelles qui définissent les F_x ainsi que celui du polynôme dont les zéros limitent les différentes régions où F_x est rationnelle (a priori k et q dépendent de x, on les suppose bornés, voire constants).

On notera que dans le cas polynomial, F est une application polynomiale de $\mathfrak{R}^{\mathfrak{p}}$ dans l'espace des fonctions bornées muni de la norme du sup.

Ainsi formulé, le problème a une généralisation évidente à l'approximation polynomiale d'un compact dans un espace normé quelconque : étant donné un ensemble F dans un espace normé B, on cherche les applications polynomiales de $\mathfrak{R}^{\mathfrak{p}}$ dans B dont l'image approche F à ε près. Ici comme par la suite, on dit que E approche F à ε près si tout élément de F se trouve à une distance au plus ε de E. En d'autres termes, plus formels :

$$sup_{x\in F} \cdot inf_{y\in E} \cdot d(x,y) \leq \varepsilon.$$

3 La variation d'un ensemble et les inégalités fondamentales

Une des idées clef de Vitushkin est d'établir des inégalités minorant la complexité d'un ensemble rationnel par morceaux qui approche un compact à ε près en fonction de l'ε-entropie de ce dernier. Il faut le faire d'abord dans un espace de dimension finie m qu'il faut imaginer très supérieure à p. C'est pour ce faire que la variation est l'outil essentiel.

Soient donc F un ensemble fermé de $\mathfrak{R}^{\mathfrak{m}}$ que dans un premier temps nous supposerons muni de la norme euclidienne, B une boule ouverte et k un entier au plus égal à m. Notons A_k l'ensemble des sous-espaces affines et généralement $v^0(E)$ le nombre de composantes connexes compactes de E. La fonction de multiplicité de F dans B est celle qui fait correspondre au sous-espace affine A le nombre $v^0(F \cap B \cap A)$.

Une question préalable se pose : la mesurabilité de cette fonction (pour la mesure μ_k invariante par le groupe des isométries qu'il faut d'ailleurs normaliser convenablement). Vitushkin en donne une démonstration très compliquée que je n'ai pas réussi à simplifier. Nous sommes en 1959. Tout analyste raisonnablement compétent sait que pour définir une fonction non mesurable, il faut utiliser l'axiome du choix. Il sait d'ailleurs que ce n'est pas un théorème et démontre la mesurabilité des fonctions qu'il utilise. Cela ne le gêne pas beaucoup parce que cette démonstration est standard, la fonction de multiplicité est la seule exception que je connaisse.

La situation a évolué depuis avec le théorème d'indépendance de Solovay [8]. Si les axiomes de Zermelo-Fraenkel ne sont pas contradictoires avec l'axiome du choix, ils ne le sont pas non plus si on les complète par les deux suivants :

(i) Tout sous ensemble de $\mathfrak{R}$ est mesurable au sens de Lebesgue

(ii) L'axiome des choix dépendants : si une relation binaire est telle que tout élément de son domaine est lié au moins à un autre, il existe une suite d'éléments chacun lié au suivant.

Le deuxième, qui est une forme très affaiblie de l'axiome du choix, permet de faire la théorie de la mesure. En particulier, on peut démontrer dans ce système que tout sous-ensemble d'une variété est mesurable pour les mesures à densité. On se ramène banalement au cas d'un pavé et on peut alors remarquer que l'image de la mesure de Lebesgue sur $[0,1]$ par l'application de Peano est la mesure de Lebesgue sur $[0,1]^m$.

Cependant on ne peut utiliser les résultats ainsi obtenus que si on est sûr de ne pas utiliser l'axiome du choix par ailleurs (par exemple, pour les analystes, Hahn-Banach dans un espace non séparable ou une conséquence du théorème de Tykhonov). Dans Zermelo-Fraenkel avec choix, il ne me semble pas que la collection des ensembles définissables sans l'axiome du choix soit un ensemble ce qui fait que l'énoncé "pour définir une fonction non mesurable il faut utiliser l'axiome du choix" ne peut pas être un théorème. Une autre difficulté vient de la possibilité que la mesurabilité d'un ensemble dûment construit dans Zermelo-Fraenkel avec l'axiome des choix dépendants (ii) soit indécidable dans Zermelo-Fraenkel avec choix. Et pourtant, compte tenu du théorème de Solovay, un tel énoncé devient très intuitif. Un résultat de ce genre serait d'autant plus intéressant que Vitushkin (qui a manifestement d'autres applications en vue) signale qu'il ne sait pas démontrer la mesurabilité pour d'autres ensembles que les fermés. [10] traite aussi les ensembles analytiques pourvus d'une propriété supplémentaire.)

Revenons à notre fonction de multiplicité. Quoi qu'il en soit, elle est mesurable (et même borelienne). On définit la variation en dimension k de F dans B comme :

$$v^k(F;B) = \int_{\mathcal{A}_{n-k}} v^0(F \cap B \cap A) d\mu_{n-k}(A),$$

que cette intégrale soit finie ou pas. La variation totale est la somme des variations dans les diverses dimensions et la dimension variationnelle $v(F;B)$ la plus grande dimension en laquelle la variation est non nulle.

Explicitons les cas les plus simples. La variation en dimension zéro n'est autre que le nombre de composantes connexes compactes. En dimension m c'est tout simplement la mesure de Lebesgue (éventuellement à un coefficient près). En dimension un, la variation d'une courbe rectifiable est sa longueur [1]. La normalisation consiste évidemment ici à donner la mesure un à l'ensemble des hyperplans qui coupent un segment de longueur un. Le résultat a été généralisé depuis ([16] et d'autres).

La définition que je viens de donner est celle que Vitushkin donne dans son livre de 1955. Cependant, dans le problème des tables, il ne s'intéresse pas à la norme euclidienne mais au maximum des valeurs absolues des coordonnées. Comme il apparaîtra mieux dans la troisième partie, c'est nécessaire vu la nature du problème qu'il traite. Il ne travaille alors qu'avec les sous-espaces parallèles à un certain nombre de vecteurs de base et prend

la moyenne des C_n^k, intégrales obtenues pour la variation en dimension k. C'est la source d'importantes complications formelles. Les boules sont désormais celles de la norme du sup.

Une propriété résultant immédiatement de la définition nous sera nécessaire par la suite :

Soient B_j des boules deux à deux disjointes contenues dans B. On a :

$$v_k(F;B) \geq \sum v_k(F;B_j) \tag{1.27}$$

En particulier :

$$v^k(F;B) \text{ croît avec } B$$

(c'est le résultat précédent dans le cas d'une seule boule B_1).

L'étape suivante consiste à majorer la variation d'un ensemble rationnel par morceaux (pour comprendre la méthode il est conseillé de penser seulement à l'image de $\mathfrak{R}^p$ par une application rationnelle). On dispose à cet effet d'une inégalité due à Oleinik majorant la somme des nombres de Betti, a fortiori le nombre de composantes connexes, d'une hypersurface algébrique réelle lisse. Ce résultat a été complété, bien après le livre de Vitushkin, par Milnor et Thom qui en particulier se passent de la régularité. Dans tous les cas l'ingrédient essentiel des démonstrations est la théorie de Morse.

Je passe à ce qui me paraît la propriété la plus profonde de la variation et qui est sans conteste la plus difficile à démontrer. On appelle profondeur d'un ensemble fermé F dans la boule B le maximum de la distance à la frontière de la boule des points de $F \cap B$. Des inégalités minorent la variation de F dans B en fonction de sa profondeur. Voici la plus simple, à défaut d'être la plus fine :

Soit F un ensemble fermé de profondeur $\rho \leq 1$ dans la boule B et de dimension variationnelle p. On a :

$$v(F;B) \geq \frac{\rho^p}{2^n} \tag{1.28}$$

Ces propriétés permettent de comprendre le rapport entre variation et approximation d'un ensemble. Revenons dans $\mathfrak{R}^n$. Soient F et E deux sous-ensembles d'un espace métrique tels que E approche F à ε près. Soit B une boule de rayon r et supposons pour simplifier que F soit contenu dans la boule concentrique de rayon $r - 2\varepsilon$. Soient $\{x_1, ..., x_N\}$ un ensemble maximal de points de F distants deux à deux d'au moins 4ε. L'ensemble E a au moins un point dans chaque boule de centre x_j et de rayon 2ε et par conséquent, sa profondeur dans la boule Bj de rayon 2ε et de centre x_j est au moins égale à ε. On a donc la chaîne d'inégalités :

$$v(E;B) \geq \sum_{j=1}^{N} v(E;B_j) \geq N\frac{\varepsilon^p}{2^n} \geq \varepsilon^p 2^{H_{4\varepsilon}(F)-n}$$

où p est la dimension variationnelle de E. On a utilisé successivement (1), (2) et la définition de l'ε -entropie.

4 Le problème du passage à la dimension infinie

Deux difficultés ont été laissées en suspens. D'abord le passage de $\mathfrak{R}^n$ à un espace fonctionnel qui est de dimension infinie. On s'attend à ce que ça puisse se faire puisqu'un compact peut être approché d'aussi près qu'on veut par un sous-espace vectoriel de dimension finie. Encore faut-il maîtriser la majoration de cette dimension. La deuxième difficulté est plus profonde, c'est que jusqu'ici on a considéré une seule application rationnelle par morceaux, alors qu'avec la définition des tables et de leur déchiffrement que nous avons donnée, cette application dépend de la variable x de façon absolument quelconque.

Vitushkin résout l'une et l'autre difficultés en introduisant un nouvel objet, qu'il ne nomme d'ailleurs pas. C'est une version abstraite de la table lue sans interpolation que j'ai donnée comme exemple dans la première partie de cet exposé. Cependant, il faut qu'il soit bien clair qu'il ne sert pas ici à construire des tables mais comme auxiliaire technique (important). Soient G un compact métrique et F un ensemble compact de fonctions continues sur G. A un sous-ensemble fini α à m points x_j de G on fait correspondre l'application Φ de F dans $\mathfrak{R}^m$ qui à la fonction f fait correspondre $(f(x_1), ..., f(x_m))$. Considérons une table quelconque, de précision ε, à p paramètres et dont l'algorithme de déchiffrement est défini par F'_x. L'application qui à $y \in \mathfrak{R}^p$ fait correspondre $(F'_{x1}(y), ..., F'_{xm}(y))$ a les deux propriétés suivantes :

- c'est une fonction rationnelle par morceaux avec les mêmes paramètres de complexité p, k, q que les F'_x ;
- son image approche $\Phi(F)$ à ε près.

Ces propriétés permettent d'appliquer les résultats obtenus en dimension finie.

Encore faut-il contrôler m ce qui n'est pas une mince affaire. Vitushkin y parvient pour deux classes de familles de fonctions. Il appelle l'une d'elles les C-espaces (je n'ai pas compris pourquoi). Elle est définie par l'existence de deux constantes A et B telles que pour tout ε strictement positif, on peut trouver un ensemble a tel que $\Phi(F)$ contienne un cube de rayon ε et que soit vérifiée l'inégalité :

$$m \geq A\, H_{B\varepsilon}(F).$$

Cette classe comprend l'ensemble des fonctions définies sur un compact métrique quelconque et admettant un module de continuité donné. Elle comprend aussi l'ensemble des fonctions s fois dérivables dont les dérivées d'ordre s vérifient une condition de Hölder.

L'autre classe, que Vitushkin appelle celle des familles facilement représentables, a des propriétés en quelque sorte opposées. Le théorème d'Arzelà montre qu'il est toujours possible de trouver un ensemble α tel que $H_\varepsilon(\Phi(F)) \geq H_{\varepsilon'}(F)$ pourvu que ε' soit strictement plus grand que ε (c'est un peu plus compliqué pour les fonctions à valeurs complexes). Sans entrer dans les détails, disons que la famille est facilement représentable si on peut régler le rapport entre ε et ε' de telle façon que m tende vers l'infini beaucoup moins vite que $H_{\varepsilon'}(F)$ quand ε et ε' tendent vers zéro. Les familles usuelles de fonctions analytiques réelles et complexes sont facilement représentables.

5 En guise de conclusion

On propose au lecteur qui aurait suivi jusqu'ici un exercice de contrôle. A supposer que les *Mathematical Reviews* lui demandent un résumé du livre de Vitushkin, dans quelles rubriques de la classification le mettrait-il ? Il doit y en avoir dans les 4000, mais même en s'en tenant aux grandes rubriques, on a le choix entre complexité, géométrie intégrale, géométrie des espaces de Banach, théorie de l'approximation, et j'en oublie sans doute. La réponse dépendra sans doute très lourdement du goût mathématique de chacun. En tout cas c'est une tâche presque impossible de trouver dans les *Mathematical Reviews* la postérité de l'ouvrage. Je soupçonne qu'il en ait eu fort peu, faute peut-être d'une communauté (est-ce le mot ?) de spécialistes pour l'accueillir.

Bibliographie

[1] J. Favard, *Une définition de la longueur et de l'aire* C.R.A.S. Paris 194 344-346, 1932.

[2] L.D. Ivanov , *Variatsij Mnozhestv i Funktsij* Nauka (Moscou), 1975.

[3] G. Lorentz, *Metric Entropy and approximation* Bull. Am. Math. Soc. 72 903-937, 1966.

[4] G. Lorentz, *Approximation of Functions* Holt, Rinehart and Winston (New York), 1966.

[5] J. Milnor, *On the Betti numbers of real varieties* Proc. Am. Soc. 15,2 275-280, 1964.

[6] O. A. Olejnik, *Ocenka czisel Betti dejstvitel'nykh algebraiczeskikh poverkhnostej* Mat. Sbornik 28(70),3, 635-640, 1951.

[7] R. Rashed, *El-Samaw'al, El-Bī rūnī et Brahmagupta : les méthodes d'interpolation* Arabic Science and Philosophy 1, 101-160, 1991.

[8] R. M. Solovay, *A model of set theory in which every set of reals is Lebesgue measurable* Ann. of Math. 92, 1-56, 1970.

[9] R. Thom, *Sur l'homologie des variétés algébriques réelles in Differential and combinatorial Topology. A Symposium in Honour of Marston Morse* Princeton University Press (Princeton) 255-265, 1965.

[10] A. G. Vitushkin, *O mnogomernykh Variaciakh* Gos. izd. tekhniko-teoreticheskoj literat. (Moscou), 1955.

[11] A. G. Vitushkin, *Otsenka Slozhnosti Zadatchi Tabulirovaniya (Evaluation de la Complexité du Problème de la Tabulation)*, Moscou, Edition d'Etat de littérature physico-mathématique, 1959.

[12] A. G. Vitushkin, *Theory of the Transmission and Processing of Information* Pergamon press (Oxford), 1961.

[13] H. Warren, *Partition by real algebraic varieties and applications to questions of non linear approximation* Bull. Am. Math. Soc. 73, 167-178, 1967.

[14] H. Warren, *Lower bounds for approximation by non linear manifolds* Trans. Am. Math. soc. 133, 192-194, 1968.

[15] M. Zerner, *Approximation par des variétés algébriques dans les espaces hilbertiens (Variation sur un thème de Vitushkin)* Studia Mathematica 38, 413-435, 1970.

[16] V. A. Zhil'tsov, *Dokazatel'stv soizmerimosti mer Favara i Khausdorfa dlja mnozhestv s ograniczennymi variatsjami* Izvzstja Ak. Nauk SSSR 30, 1157-1162, 1966.

View publication stats

6 Annexe 1 : liste des chapitres

VITUSHKIN A. G. 1959 *Otsenka Slojnosti Zadatchi Tabulirovaniya (Evaluation de la Complexité du Problème de la Tabulation)* , Moscou, Edition d'Etat de littérature physico-mathématique

I Le concept d'entropie d'un espace métrique
II L'entropie des espaces de fonctions analytiques
III L'entropie de quelques sous-espaces des fonctions continues
IV Les variations d'un ensemble
V Evaluations de variation pour quelques ensembles particuliers
VI Familles de fonctions facilement représentables
VII Représentations de quelques classes de fonctions continues

Annexe 2 : les domaines visités

Mesurabilité et logique
Géométrie intégrale
Géométrie algébrique réelle
Topologie algébrique
Théorie de Morse